DOCUMENT RESUME

ED 386 042 FL 023 205

AUTHOR	Bird, Norman, Ed.; And Others
TITLE	Language and Learning.
INSTITUTION	Hong Kong Education Dept. Inst. of Language in Education.
REPORT NO	ISBN-962-7854-38-7
PUB DATE	94
NOTE	562p.; Selected papers presented at the Annual International Language in Education Conference (Hong Kong, 1993). For individual papers, see FL 023 205, FL 023 207-223, FL 023 225-233, and ED 366 228, ED 368 181, ED 370 411, ED 371 618, and ED 372 600.
PUB TYPE	Collected Works - General (020) -- Speeches/Conference Papers (150)
EDRS PRICE	MF02/PC23 Plus Postage.
DESCRIPTORS	*Chinese; Classroom Techniques; *Cognitive Processes; Educational Strategies; Educational Television; Elementary Secondary Education; *English (Second Language); Foreign Countries; *Instructional Materials; *Language Acquisition; Language Patterns; Language Processing; Language Research; Language Teachers; Language Tests; Linguistic Theory; Oral Language; Orthographic Symbols; Reading Comprehension; Reading Instruction; Second Language Instruction; Second Language Learning; *Second Languages; Teacher Education; Testing; Verbal Tests; Vocabulary Development; Writing Instruction; Written Language
IDENTIFIERS	Authentic Materials; *Hong Kong; Japan

ABSTRACT

Papers from a 1993 International Conference on Language in Education include: "A Language Development Approach to Education" (M. A. K. Halliday); Text, Talk, and Inquiry: Schooling as a Semiotic Apprenticeship" (G. Wells); "Chinese Orthography and Reading" (O. J. L. Tzeng); "Task-Centred Assessment in Language Learning" (G. Brindley); "Task as a Unit of Teaching Analysis" (S. J. Hall); "Using a Functional Approach in Assessing Written Texts" (D. Nunan); "Steppe by Step: A Cultural Approach to Language Tasks" (C. Barron); "Coherence and Continuity in the Task-Centred Language Curriculum: Global Education as a Framework for Task-Based Language Teaching" (B. Bushell, B. Dyer); "From Task Description to Task Enactment: Teachers' Interpretation of Language Learning Tasks" (G. T. Sachs, S. Kong, A. Lo, T. Lee); "How Students Revise Propositions" (P. Falvey, S. Sengupta); "Writing from Sources: Does Source Material Help or Hinder Students' Performance?" (J. A. Lewkowicz); "Effects of Signalling on Reading Comprehension" (I. K. B. Lee); "The Variety and Effectiveness of Strategies Employed in Vocabulary Explanations in EFL Classes in Hong Kong" (A. L. On-lai); "English Extensive Reading in the Primary Curriculum" (V. Yu, E. Chu, S. Yuen-lan, R. Yau); "Investigating Lexis beyond the Most Frequent Words, Part 2" (N. Bird); "The Future Role of EFL Textbook Resources in Hong Kong" (B. Adamson, J. C. K. Lee); "Concordancing for School" (V. Pickard, K. Chan, J. Tibbetts); "What Makes Authentic Materials Different? The Case of English Language Materials for Educational Television" (A. McNeill); "Resources for the Independent EFL Learner in Japan" (S. M. Ryan); "Action Research Contributes More to Teaching Than Just Solving Discrete Problems in the Classroom" (A. Mueller); "Computer-Mediated Communication and Teacher Education: The Case of Telenex" (D. Coniam, S. Sengupta, A. B. M. Tsui, W. Kam-yin); "Vocabulary Learning Strategies of Good and Poor Chinese EFL Learners" (G. Yong-qi); "Chinese Readers' Metacognitive Awareness in Reading Chinese and English" (C. Chern); "The Impact of Illustrations and Cultural Schemata on Hong Kong Pupils' Reading Comprehension and Recall of Text" (T. Dolan); "Word Printed Frequency/Familiarity and Structure Complexity Effects on L1 and L2 Word Recognition Processes in Chinese" (Y. Sun); "A Word in Your Ear: To What Extent Does Hearing a New Word Help Learners To Remember It?" (M. Hill); "Pig in the Middle? Effects of Mediating Tasks on Cognitive Processing of Text" (D. M. Allison, V. Berry, J. Lewkowicz); "The Assessment of Spoken Language Under varying Interactional Conditions" (V. Berry); "The Grammatical Awareness and Knowledge of Hong Kong Teachers of English" (S. Andrews); "Some Characteristics of Native and Non-Native Speaker Teachers of English" (A. McNeill); and "Using Concurrent Verbal Reports as Teaching Tools in Language Teacher Education" (M. Falvey). (MSE)

LANGUAGE AND LEARNING

ED 386 042

Edited by
Norman Bird
with
Peter Falvey
Amy B. M. Tsui
Desmond M. Allison
Arthur McNeill

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. L. Clark

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL023205

Institute of Language in Education
Education Department

Department of Curriculum Studies
The University of Hong Kong

Hong Kong Association of Applied Linguistics

BEST COPY AVAILABLE

LANGUAGE AND LEARNING

Edited by

Norman Bird (Editor-in-Chief)

Peter Falvey
Amy B.M. Tsui
Desmond M. Allison
Arthur McNeill

Technical editor and production manager

Roger J. Webb

Institute of Language in Education

Education Department

Hong Kong

1994

Copyright (c) 1994, Institute of Language in Education,
Hong Kong Education Department.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photographic, or otherwise, without the prior permission of the Editor.

The articles in this book record the personal views of the contributors and should not be taken as expressing the official views of the Education Department, Hong Kong.

ISBN 962 7854 38 7

Cover design by Roger J. Webb

CONTENTS

LIST OF CONTRIBUTORS vi

PREFACE AND OVERVIEW x

ACKNOWLEDGEMENTS xv

PART ONE

THE PLENARY PAPERS

INTRODUCTION 3

1. A Language Development Approach to Education — M.A.K. Halliday 5

2. Text,Talk and Inquiry: Schooling as a Semiotic Apprenticeship — Gordon Wells 18

3. Chinese Orthography and Reading: A Clarification — Ovid J.L. Tzeng 52

4. Task-Centred Assessment in Language Learning: The Promise and the Challenge — Geoff Brindley 73

PART TWO

TASK-CENTRED LEARNING AND ASSESSMENT

INTRODUCTION 97

5. Task as a Unit of Teaching Analysis — Stephen J. Hall 98

6. Using a Functional Approach in Assessing Written Texts — David Nunan 126

7. Steppe by Step: A Cultural Approach to Language Tasks — Colin Barron 139

8. Coherence and Continuity in the Task-Centered Language Curriculum: Global Education as a Framework for Task-Based Language Teaching 154
Brenda Bushell and Brenda Dyer

9. From Task Description to Task Enactment: Teachers' Interpretation of Language Learning Tasks 172
Gertrude Tinker Sachs, Stella Kong, Anne Lo and Tom Lee

PART THREE

TEXT AND TALK IN LEARNING

INTRODUCTION 191

10. How Students Revise Propositions: An Exploratory Study of Propositional Modification 193
Peter Falvey and Sima Sengupta

11. Writing from Sources: Does Source Material Help or Hinder Students' Performance? 204
Jo A. Lewkowicz

12. Effects of Signalling on Reading Comprehension 218
Icy K.B. Lee

13. The Variety and Effectiveness of Strategies Employed in Vocabulary Explanations in EFL Classes in Hong Kong 237
Annie Lee On-lai

PART FOUR

RESOURCES FOR TEACHERS AND LEARNERS

INTRODUCTION 255

14. English Extensive Reading in the Primary Curriculum: Current Practices and New Initiatives 258
Vivienne Yu, Emily Chu, Siu Yuen-lan and Rachel Yau

15. Investigating Lexis Beyond the Most Frequent Words - Part 2 274
Norman Bird

16. The Future Role of EFL Textbook Resources in Hong Kong 288
Bob Adamson and John C.K. Lee

17. Concordancing for School: Problems and Potential — 300
Valerie Pickard, Kenneth Chan and Janice Tibbetts

18. What Makes Authentic Materials Different? The Case of English Language Materials for Educational Television — 314
Arthur McNeill

19 Resources for the Independent EFL Learner in Japan — 327
Stephen M. Ryan

20. Action Research Contributes More to Teaching Than Just Solving Discrete Problems in the Classroom — 337
Annie Mueller

21. Computer-Mediated Communication and Teacher Education: The Case of Telenex (Colloquium Report) — 352
David Coniam, Sima Sengupta, Amy B.M. Tsui and Wu Kam-yin

PART FIVE

COGNITIVE PROCESSES IN LANGUAGE LEARNING

INTRODUCTION — 373

22. Vocabulary Learning Strategies of Good and Poor Chinese EFL Learners — 376
Gu Yong-qi

23. Chinese Readers' Metacognitive Awareness in Reading Chinese and English — 402
Chiou-lan Chern

24. The Impact of Illustrations and Cultural Schemata on Hong Kong Pupils' Reading Comprehension and Recall of Text — 422
Terry Dolan

25. Word Printed Frequency/Familiarity and Structure Complexity Effects on L1 and L2 Word Recognition Processes in Chinese — 433
Yilin Sun

26. A Word in your Ear: To What Extent Does Hearing a New Word Help Learners to Remember It? — 447
Monica Hill

27. Pig in the Middle? Effects of Mediating Tasks on Cognitive Processing of Text 463
Desmond M. Allison, Vivien Berry and Jo Lewkowicz

28. The Assessment of Spoken Language Under Varying Interactional Conditions 491
Vivien Berry

29. The Grammatical Awareness and Knowledge of Hong Kong Teachers of English 508
Stephen Andrews

30. Some Characteristics of Native and Non-Native Speaker Teachers of English 521
Arthur McNeill

31. Using Concurrent Verbal Reports as Teaching Tools in Language Teacher Education 533
Margaret Falvey

LIST OF CONTRIBUTORS

Bob Adamson	Sir Robert Black College of Education, Hong Kong.
Desmond M. Allison	English Centre, University of Hong Kong.
Stephen Andrews	Curriculum Studies Dept., University of Hong Kong.
Colin Barron	English Centre, University of Hong Kong.
Vivien Berry	English Centre, University of Hong Kong.
Norman Bird	Institute of Language in Education, Hong Kong.
Geoff Brindley	National Centre for English Language Teaching, Macquarie University, Sydney.
Brenda Bushell	International Christian University, Tokyo, Japan
Kenneth Chan	Language Centre, University of Science and Technology, Hong Kong.
Chern Chiou-lan	Tunghai University, Taiwan.
Emily Chiu	Institute of Language in Education, Hong Kong.
John L. Clark	Institute of Language in Education, Hong Kong.
David Coniam	TELEC, Faculty of Education, University of Hong Kong.
Terry Dolan	University of Nottingham, England.
Brenda Dyer	International Christian University, Tokyo, Japan.
Margeret Falvey	Dept. of Curriculum and Instruction, Chinese University, Hong Kong.
Peter Falvey	Curriculum Studies Dept. University of Hong Kong.
Gu Yong-qi	Education Dept., University of Hong Kong.
Stephen J. Hall	SEAMEO - RELC. Singapore.

M.A.K. Halliday	5 Laing Avenue, Kilara, NSW 2071, Australia.
Monica Hill	English Centre, University of Hong Kong.
Albert Ip	Curriculum Studies Dept., University of Hong Kong.
W.W. Ki	Curriculum Studies Dept:, University of Hong Kong.
Stella Kong	Institute of Language in Education, Hong Kong.
Lai Kwok-wing	Dept. of Education, University of Otago, New Zealand.
Annie Lee On-lai	St. Margaret's Girls' College, Hong Kong.
Icy K.B. Lee	Institute of Language in Education, Hong Kong.
John Lee	Sir Robert Black College of Education, Hong Kong.
Jo Lewkowicz	English Centre, University of Hong kong.
Anne Lo	Institute of Language in Education, Hong Kong
Arthur McNeill	Curriculum Studies Dept., University of Hong Kong.
Annie Mueller	English Centre, University of Hong Kong.
David Nunan	English Centre, University of Hong Kong.
Valerie Pickard	English Centre, University of Hong Kong.
Stephen Ryan	Dept. of General Education, Osaka Institute of Technology, Osaka, Japan.
Sima Sengupta	Curriculum Studies Dept., University of Hong Kong.
Siu Yuen-lan	Institute of Language in Education, Hong Kong.
Yilin Sun	OISE Employment Canada, Toronto Board of Education, Toronto, Canada.
Janice Tibbetts	King's College, Hong Kong.

Gertrude Tinker Sachs	Institute of Language in Education, Hong Kong.
Amy B.M. Tsui	Curriculum Studies Dept., University of Hong Kong.
Ovid J.L. Tzeng	National Chung Cheng University, Ming Hsiung, Taiwan.
Gordon Wells	Ontario Institute for Studies in Education, Canada.
Wu Kam-yin	TELEC, Curriculum Studies Dept., University of Hong Kong.
Rachel Yau	Institute of Language in Education, Hong Kong.
Vivienne Yu	Institute of Language in Education, Hong Kong.

BEST COPY AVAILABLE

PREFACE

Introduction

The International Language in Education Conference (ILEC 93) was the first joint Conference to be organised by the Institute of Language in Education, the Department of Curriculum Studies of the University of Hong Kong and the Hong Kong Association for Applied Linguistics. ILEC 93 builds on the series of annual International Conferences launched by the Institute of Language in Education in 1985. It is a deliberate attempt to widen the constituency of those directly involved in planning and organising the Conference, in the firm belief that this will make it a more representative and enriching forum for the exchange of ideas and experience among educators concerned with Chinese and English in Hong Kong and in the wider international community.

The Language in Education Context in Hong Kong

The Hong Kong Government has been particularly concerned recently to improve both academic and language proficiency levels through effective language in education policy and practices. Four major initiatives are currently being developed.

Firstly the Government has introduced a new medium of instruction policy designed to ensure at secondary school level that teachers and students hold to the consistent use of either Chinese or English for the construction and use of knowledge across the curriculum, rather than mixing and switching between the two as they do at present. This depends upon effective medium of instruction grouping procedures at entrance to secondary school, an effective bridging programme to assist those students assessed as able to learn through English to move from Chinese-medium learning to English-medium learning, should they wish to do so, and effective teacher education programmes to enable schools and teachers to develop good practices. At university level English enhancement for academic purposes has received extra funding in response to the challenge posed by the inadequate level of English of most entrants.

Secondly, the Government is developing a target-oriented curriculum, based on task-centred learning and assessment, which embraces Chinese and English as mediums of instruction and as core subjects. The emphasis is on the use of language for inquiring, constructing knowledge, logical thinking, expressing thought and solving problems across the curriculum. School assessments and examinations will require students to talk and to construct written texts to express their thinking, rather than simply tick boxes or give one word or one sentence answers. The name Target-Oriented Curriculum has now replaced the earlier name of Targets and Target-Related Assessment, since it describes the exercise we are engaged in in Hong Kong much better. It indicates that the emphasis is on the renewal of the curriculum as a whole, rather than simply on the parts concerning target-setting and assessment.

Thirdly, a Hong Kong Vocational English Programme is being developed at several levels for school leavers and employees, tailored to the needs of the workplace. It is to be certificated by the London Chamber of Commerce and Industry. There is also a need now to develop a parallel programme to meet the demand for better Chinese skills in the workplace.

Fourthly, the Government is determined to improve the quality of teacher education in Hong Kong by setting up a new tertiary level Institute of Education, through the amalgamation of the four Colleges of Education and the Institute of Language in Education. The new Institute will run upgraded courses leading to the Teaching Certificate, and alongside the Universities will start to provide both pre-service and in-service programmes leading to B.Ed degrees, and in time to PGCE, M.Ed and Ph.D awards. The Government is now committed to working towards an all graduate teaching profession. 70% of secondary teachers are graduates already, and by 2007 35% of primary teachers will also be graduates.

ILEC 93 provided those of us who work in Hong Kong in language in education with an opportunity to exchange ideas and experience with others on some of the issues arising out of the above initiatives.

ILEC 93

The theme of ILEC 93 was Language and Learning. This theme enabled us to build upon what we achieved last year in the exploration of the relationship between Language and Content in both language classrooms and content subject classrooms. Whereas last year our focus was largely on knowledge and its linguistic representation, this year the focus was on the process through which knowledge and language are developed.

We chose to focus on four sub-themes. The first sub-theme was Task-centred Learning and Assessment in language classrooms, which is of particular concern to us, since the Target-Oriented Curriculum, which is to be developed in schools in Hong Kong, places task-centred learning and assessment at the heart of the curriculum.

The second sub-theme was Text and Talk in Learning. Here the focus ranged from text and talk in the language classroom to the use of text and talk in learning in general, to different spoken and written genres, and to the interdependence between language learning and learning in general.

The third sub-theme was entitled Resources for Teachers and Learners. This enabled us to examine the range of resources learners need, and to explore ways in which teachers can best be supported in their work through the setting up of teacher resource centres and computerised networks.

The fourth sub-theme was entitled Cognitive Processes in Language Learning. This enabled us to take stock of where we are in our understanding of the process of language learning and to explore implications for the classroom.

All four themes were applied to school, tertiary, vocational and teacher education contexts.

For staff of the Institute of Language in Education ILEC 93 was an occasion for mixed emotions. It was the last conference at which the Institute would feature as an independent body within the Education Department, and at the same time it represented the first step along the road to being disestablished and joining other tertiary institutions in Hong Kong as part of the new Hong Kong Institute of Education. The new Institute is expected to be up and running in September 1994, and to move into a new purpose-built rural campus in Taipo in the New Territories in 1997. I believe that there is everything to be gained from having language in education firmly rooted once again within education across the curriculum as a whole, rather than hived off from it in an Institute of Language in Education. There is also everything to be gained from being disestablished from the civil service with its bureaucratic procedures to join the academic community in which teacher education, research and curriculum development work belong. It is my fervent hope, however, that disestablishment will be counterbalanced by the establishment of a much better structure for coordination between teacher education, curriculum development, curriculum monitoring and evaluation, and examinations.

I am particularly pleased that the ground has now been laid and the budgeting arrangements made for the continuation of ILEC in future years, with the Hong Kong Institute of Education playing a major organisational role within it, through its School of Language which will be formed by the union of the language departments in the four Colleges of Education and the Institute of Language in Education.

ILEC 94

ILEC 94 will be organised jointly by the Hong Kong Institute of Education, the Curriculum Studies Department of the University of Hong Kong, the English Centre of the University of Hong Kong and the Hong Kong Association for Applied Linguistics. For ILEC 94 it is the turn of the Curriculum Studies Department of the University of Hong Kong to take the role of lead organiser.

ILEC 94 will take place from December 14 to 16 in the University of Hong Kong. Four themes within Language in Education will be focused upon. They are:

- Implementing change in the classroom
- Mother tongue education in the Chinese context
- Teachers and language awareness
- Language and the world of world

ILEC 93 Proceedings

This book is a selection of Conference Papers presented at ILEC 93. We hope that it will be of as much value to others as it will be to those of us who work in language in education in Hong Kong.

John L. Clark
Director, Institute of Language in Education
Chairman, ILEC 93 Planning Committee

OVERVIEW

The first International Language in Education Conference (ILEC93) held in December 1993 replaces the former Institute of Language in Education Conference which had been held in the previous eight years since 1985.

As a result of this change it has been possible to call upon more expertise than was previously available both in selecting the papers for presentation and editing them for publication. Furthermore, in editing the papers I have had the pleasure of working with four editors from Hong Kong University and five sub-editors from the Institute of Language in Education.

The papers read at the 1993 Conference were in Cantonese, Putonghua and English, and the present volume contains a selection of those papers that were presented in English. This selection is organised in five parts.

The papers in Part I contain the four plenary papers.

The five papers in Part II, which were edited by Peter Falvey, address the conference's first sub-theme, *task-centred learning and assessment*.

The four Part III papers edited by Amy B.M. Tsui deal with the second sub-theme, *text and talk in learning*.

The eight papers in Part IV, which were edited by Desmond M. Allison, treat the third sub-theme, *resources for teachers and learners*.

The ten Part V papers edited by Arthur McNeill consider the fourth sub-theme, *cognitive processes in language learning*.

As more people have helped in getting this book into print this year, it has been possible to publish it within six months of the conference inside of the usual twelve. Furthermore, as all the editors are experts within their own fields, we hope that you will find a greater variety of editorial comment than has been possible in the previous conference papers.

Norman Bird
Editor-in chief, ILEC 93 Conference Proceedings.

ACKNOWLEDGEMENTS

As chief-editor with the new task of both editing and co-ordinating the editing of the first International Language in Education Conference (ILEC93) Proceedings, I would like to express on behalf of the editors our gratitude to the members of the staffs of the Institute of Language in Education and of the University of Hong and to the members of the Hong Kong Association of Applied Linguists (HAAL) for their valuable advice and assistance in the preparation and processing of this volume. I would especially like to thank, however, the editors, the proof-readers, sub-editors listed below, and the technical editor and production manager Roger Webb without whose help the whole venture would simply have been impossible. Finally, I would also like to express my particular thanks to the Hong Kong Government Printer and the members of his staff in the Printing Department for their courtesy and efficiency.

Norman Bird (Editor-in-chief)
Institute of Language in Education
2 Hospital Road
Hong Kong

Sub-editors

David Carless
John Duncan
Mandy Evans
Mike Ingham
Gertrude Tinker Sachs

Proof-readers

David Bunton
John L. Clark
Denis Donahue
John Harris
Henry Hepburn

PART ONE

THE PLENARY PAPERS

INTRODUCTION

Four plenary papers were given in English at this year's conference.

The keynote plenary speaker was *M.A.K. Halliday*, whose paper entitled *A Language Development Approach to Education* was based on the insights of a long and distinguished life in academic research. The paper plots the stages of children's linguistic (especially grammatical), psychological and educational development from the earliest years through the child's time at school. He emphasizes in particular the "discontinuity between educational and commonsense forms of knowledge," and explains both the problems of and benefits for a child in coming to terms with shifts in modes of meaning, e.g. from narrative to paradigmatic, or later from abstract to metaphorical. He then goes on to apply his theories to educational practice today, i.e. "[those] aspects of children's language development which seem to me to offer pointers to the nature of learning in general." Of particular interest to those working in education in Hong Kong today, are *Halliday's* references to patterns of regression and reconstruction at the different stages of a child's educational development, which allows the learner to create new dimensions of semantic space. If this can be borne in mind by Hong Kong's curriculum planners, a lot of frustration and anguish may be avoided among Hong Kong's population of eleven-year-olds, as, to use *Halliday's* words the children "progress from a less dynamic into a more synoptic vision of the world."

Gordon Wells in his plenary paper entitled *Text, Talk and Inquiry: Schooling as Semiotic Apprenticeship* elaborates on the implications of *Halliday's* paper in terms of educational practice. Drawing on sociocultural theory, *Wells* defines the activity of eduction as "the selection and organization of joint activities that provide the opportunity for students to appropriate the practices and artifacts valued by their culture and to use these resources in creative ways to solve new problems in tasks which they have a part in choosing and directing, and in which they receive guidance and assistance from peers and teachers." He then goes on to explain the role of language in its various genres in the curriculum, and observes that "as the most powerful and versatile of culture's resources, discourse is both used to learn and learned in use." He finally argues that such a conception of education as he proposes is best enacted through an inquiry-oriented curriculum developed by everyone involved in education at every level each taking into consideration individual schools and classes operating against the background of particular cultural and historical settings.

Ovid J.L. Tzeng's paper *Chinese Orthography and Reading: A Clarification* might be taken as a good example of *Wells'* concluding exhortation that individual educationists study the special needs of individual education systems with particular regard to individual historical and cultural circumstances. This very specific but nevertheless fascinating paper compares and contrasts the origins of Chinese logographs and western phonemic writing system from a mainly historical point of view and considers the implications for neural processing. With the aid of an impressive list of references *Tzeng* examines in particular those research findings which address the following five questions:

1. Are Chinese logographs difficult to learn?
2. Does reading Chinese require an enhanced visual memory?
3. Is it true that the skilled reading of Chinese does not require the speech recoding process?
4. Is it true that phonological awareness is not essential for learning to read the Chinese script?
5. Does reading Chinese involve greater right hemispheric brain processing than reading alphabetic scripts?

As some oriental education systems have renounced logographic writing systems in favour of alphabetic ones for their native languages in recent years, e.g. Vietnamese and Korean, and as Japanese still uses a compromise system involving both approaches, *Tzeng's* clear and succinct paper is obviously of great importance to everyone concerned with the development of education in the Orient in the immediate future.

The final plenary paper, *Geoff Brindley's Task-Centred Assessment in Language Learning: The Promise and the Challenge* considers how the numerous developments now taking place in language learning can best be assessed. *Brindley* outlines some of the attractions of task-centred assessment (TCA) such as directness and authenticity, but he also points out that, like other forms of assessment, TCAs need to demonstrate their validity, reliability and practicality. *Brindley* explores four methods of deriving assessment criteria, 'expert judgement' approaches, rating scales, genre-based approaches and data-based criteria, and notes how technological advances can aid inter-marker reliability via multi-faceted Rasch analysis. He concludes pragmatically that TCA may oblige practitioners to rethink the notion of high-agreement reliability and its implications in terms of time allocation, resourcing levels and professional development.

In conclusion, it only remains to note that the four plenary papers we are fortunate enough to publish here range over a vast field of educational research and endeavour, but like all good papers, they each have one thing in common; all the presenters consider the practical implications of their ideas for educational progress in general, and also how they will effect classroom teachers, as they go about their daily routines.

Norman Bird

A LANGUAGE DEVELOPMENT APPROACH TO EDUCATION

M.A.K. Halliday

It is a great pleasure for me to be here in Hong Kong on this occasion, and to be taking part in the International Language in Education Conference "ILEC 93". The theme for this year's conference is "Language and Learning", and I have tried to locate my own contribution squarely within that topic. For me the conference comes towards the end of a few weeks' stay in Hong Kong, during which I have been working with colleagues in the language education area; and one of the issues that we have been exploring is that of the relation between commonsense learning and educational learning - between the kind of learning that children are involved in, more or less from birth, in the family and among their own peer group, and the kind of learning they engage in when they come to school, where learning is institutionalized (that is, after all, what a school is: an institution designed for learning in). These two aspects of children's learning experience, commonsense learning and educational learning,.are not of course insulated one from the other: there is continuity between the two; but there is not perhaps as much continuity as there could be, and some people might feel that the two are kept rather too far apart. In Hong Kong this is probably thought of as a consequence of the language situation, given the distance that typically separates the language of home from the language of school. This obviously plays some part. But lack of continuity between commonsense and educational learning is not just a feature of societies that are linguistically complex. Even where home and school share essentially the same language of interaction, there is typically a considerable discontinuity in children's experience of learning, as they move between these two learning environments.

Now this is not the principal focus of my talk today; but I need to look a little further into the phenomenon of learning discontinuity, in order then to look behind it and beyond it. What is the nature of this discontinuity between home and school, and how does it arise? One factor is presumably the linguistic medium: commonsense learning, in the pre-school years at least, is thoroughly grounded in the spoken language; whereas after children become literate, at the very beginning of their stay in school, it is typically assumed that what they learn in class will be learnt essentially through reading and writing. But this is clearly not the whole of the picture. After all, even in school the teacher talks to them, and they discuss what they are learning both with the teacher and with each other; and on the other hand, before children ever go into school their parents are often reading to them out of books, and some children learn to read quite a lot all by themselves. So there is no exact equation such that commonsense learning equals learning through speech and educational learning equals learning through writing. Nevertheless the difference between speech and writing is a significant factor - although we should concentrate, rather, not on the medium itself but on the difference between spoken language and written language. It is not the different media that are relevant so much as the different kinds of meaning that are typically associated with them.

What we are observing, in this context, is a discontinuity between educational and commonsense forms of knowledge: between two different ways of construing human experience. It is obviously impossible to characterize this difference adequately in a few short sentences; it is something complex and many-sided. But I can try and capture one or two salient points. (1) Commonsense knowledge is fluid and indeterminate, without clear boundaries or precise definitions: it does not matter too much exactly where a particular process begins and ends, or what is one phenomenon and what is another. Educational knowledge is determinate and systematic: the categories of experience are organized into conceptual structures with defined properties and explicit interrelations. (2) Commonsense learning foregrounds processes - actions and events, including mental and verbal events; of course it is also concerned with things, but their main significance is in the way they enter in to all the various processes. Educational knowledge foregrounds the things: persons and concrete objects, then later on increasingly abstract and virtual objects that are needed to explain how the things behave. (3) Commonsense knowledge is typically construed as dialogue, and built up interactively, or "intersubjectively", by the human group. Educational knowledge is typically construed monologically, and built up by each individual - the "others", in our present educational system at least, tend to be competitors rather than collaborators. (4) And commonsense knowledge is typically unconscious: we do not know what we know; whereas educational knowledge is conscious knowledge - and so it can be rehearsed, and therefore monitored and assessed. There are no examinations for knowledge of the commonsense kind.

James Britton, in his influential book *Language and Learning* written about a generation ago, distinguished in students' writing between the private, "expressive" kind and the more public kinds demanded by the school, "transactional" on the one hand and "poetic" on the other. Britton saw the expressive as the learner's point of departure, the natural mode of meaning that children brought with them from the experience of their early years. The priority that Britton gave to the expressive category derived from his own rather individualistic ideology of education; but his work had considerable influence on educational practice in England and elsewhere - for example in the way primary school writing came to be dominated by stories, on the assumption that the bridge from commonsense to educational learning was to be built out of personal narrative. (See Britton, 1970.) Narrative is, in turn, the term that Jerome Bruner uses to name one of his two modes of "cognitive functioning", the narrative and the paradigmatic. The paradigmatic mode "attempts to fulfil the ideal of a formal, mathematical system of description and explanation: It employs categorization or conceptualization and the operations by which categories are established, instantiated, idealized, and related one to the other to form a system." By contrast, "the imaginative application of the narrative mode leads instead to good stories, gripping drama, believable (though not necessarily "true") historical accounts. It deals in human or human-like intention and action and the vicissitudes and consequences that mark their course." These two modes of cognitive functioning each provide, according to Bruner, "distinctive ways of ordering experience, of constructing reality." (See Bruner 1990:11-13.)

We see this dichotomy transformed and built into educational knowledge if we compare the language of natural science and the language of the humanities, as Martin and his colleagues have done in their detailed studies of these discourses in the secondary school (see Halliday & Martin 1993: esp. chapter 11). The grammar of science constructs elaborate technical taxonomies, using nominalizing metaphors and complex nominal group structures to create virtual objects and build them into sequences of logical argument. The grammar of the humanities, on the other hand, constructs schemata made up of individual semi-technical abstractions, simpler in structure (often single nouns) because not taxonomized, but each one charged with value and coming together as a whole to make up an ideological stance. Compare the following two passages, the first taken from a geography textbook and the second from a textbook of history:

> *As air is moved upward* away from the land-water surface or downward towards it, very important *changes occur in the air temperature. Air moving upward* away from the surface comes under lower pressures because there is less weight of atmosphere upon it, so *it stretches or expands.* Air moving downward towards the surface from higher elevations encounters higher pressures and *shrinks in volume.* Even when *there is no addition or withdrawal of heat* from surrounding sources, the temperature of *the upward or downward-moving air* changes because of *its expansion or contraction. This type of temperature change* which results from *internal processes* alone is called *adiabatic change.*
>
> [G.T. Trewartha, *An Introduction to Climate,* 1968:1361]

I have used italics to mark examples of how the grammar constructs technical entities and organizes them into logical sequences; e.g. *[air] stretches or expands ... because of its expansion or contraction; changes occur in the air temperature ... this type of temperature change ... is called adiabatic change.*

> Wars are costly *exercises.* They cause *death* and *destruction* and *put resources to non-productive uses* but they also *promote industrial and technological change.* This *benefit* does not mean that war is a good thing, but that sometimes it *brings useful developments.*
>
> The Second World War further *encouraged the restructuring of the Australian economy* towards *a manufacturing basis.* Between 1937 and 1945 the value of *industrial production* almost doubled. *This increase* was faster than otherwise *would have occurred. The momentum was maintained* in the post-war years and by 1954-5 *the value of manufacturing output was three times* that of 1944-5. *The enlargement* of Australia's *steel-making capacity,* and of chemicals, rubber, metal goods and motor vehicles all *owed* something to the *demands* of war. The war had acted as a hot-house for *technological progress* and *economic change.*
>
> [H. Simmelhaig and G.F.R. Spencely, *For Australia's Sake,* 1984:121]

Here the italics show instances of abstract expressions of a semi-technical kind (e.g. *exercises, put ... to non-productive uses, brings ... useful developments*) and terms with a clear evaluative loading (e.g. *destruction, non-productive, promote, benefit, useful, increase, momentum*). The ideological motif of 'growth is good' is foregrounded throughout (cf. Halliday 1993a:25 ff.).

I will refer again to these examples later on. Here the point I am drawing attention to is this: the kind of variation that we find here at secondary level, between the discourses of science and the humanities, is an elaboration of the same dichotomy; this dual motif runs throughout the educational process, and there seems no reason to assign priority to one variant or the other. Yet in much of contemporary learning theory and educational practice in the West it is assumed that the narrative mode (in Bruner's sense) is somehow cognitively prior, and that commonsense learning is overwhelmingly in terms of "good stories". Bruner himself acknowledges (p. 127) that his own early model of the child was "very much in the tradition of the solo child mastering the world by representing it to himself in his own terms"; and this model readily lends itself to (and in practice typically co-occurs with) a "story-telling" interpretation of childhood. I think that we, as educators, should challenge and be prepared to reject this kind of "childist" model. If we accept any such dichotomy as that proposed by Bruner (and it may be helpful as a tool for thinking with, although we might adapt it to become less dichotomized and more explicitly grounded in language), we probably need to recognize that both these modes of meaning, the paradigmatic as well as the narrative, contribute equally to children's commonsense ordering of experience.

If we are seeking a model from educational theory which we can relate to the distinction between commonsense and educational knowledge as this is manifested in children's early language development - where the commonsense reality is construed in language before the educational one - we might do well to reexamine Bernstein's theory of code, deriving from a sociological rather than a psychological perspective on learning. Commonsense and educational learning construe reality in terms of different codes. While these do not correspond exactly to Bernstein's "restricted" and "elaborated" varieties (there can be various features of elaborated code in the linguistic construction of commonsense knowledge), they are related at a general level; and more specifically, in that educational knowledge as at present constituted cannot be construed without the semantic resources that Bernstein identified as "elaborated". This applies equally both to the discourse of science and to that of the humanities.

What we have been lacking, however, it seems to me, is a perspective on learning that starts from language itself, instead of first being formulated from outside language and then mapped on to observations of language as an afterthought. Of course we have moved some way from the views of Piaget, who saw language as essentially a means for the expression of thought processes. Both Bernstein and Bruner, arguing for a constructivist view (and citing Vygotsky as a pivotal figure), foreground language as a central factor in the process by which reality is constructed. But if reality is constructed in language - or, as I would prefer to put it, if human experience is construed in the form of language - then the way in which language itself comes into being must give us an insight into the fundamental nature of learning. After all, children are at the same time both learning language and using language to learn with (as Gordon Wells has documented very richly in the course of his work). It is we who distinguish these two processes, as we have to do for purposes of analysis; as far as the children themselves are concerned, learning language and learning through language are just one integrated process - namely, learning. Might we not take more account of what has been found out about children's language development, when we try to increase our understanding of the nature of learning in general?

It seems to me that there are certain aspects of what we know about language development in children, if we start from the earliest phase before they move into the mother tongue, that are relevant and suggestive in such a context. I am not going to try to enumerate them all - I have written about this elsewhere (Halliday, 1993b); but I should like to discuss one or two of these features of children's learning which I think are particularly relevant to the present situation here in Hong Kong. Let me refer first of all to the very general principle of linguistic function, and ask: what are the functional contexts in which language first appears?

1. Very early in life, children find that they can use language - not yet the mother tongue, but a "child tongue", a little protolanguage they construct for themselves, in interacting with parents and others - in a number of different ways: to get things done for them, or given to them; to get others to join in some activity, or else just to attend to them and "be together"; and to express their own feelings and curiosity about the outside world. When they start to learn the mother tongue, however, and thus get ready to construe their experience in the distinctively human mode, children typically adopt a simple but very powerful strategy: they reconstrue these functions by setting up a very general opposition - that between language to act with and language to think with. In this period, round about the second half of the second year of life, it has often been observed that children's utterances are of one or other of these two kinds: either pragmatic - they want something done for them; or what I called "mathetic", meaning by this the learning function - they are learning to name things and to describe what is going on around them. This strategy then turns out to be a transitional one leading to something much more pervasive and lasting: before very long each utterance comes to include a combination of both functions, having both an active and a reflective dimension of meaning. Now, from the language point of view, what we are seeing here is the birth of grammar, as (i) the opposition between pragmatic and mathetic evolves into the mood system (indicative /imperative and so on), while (ii) the experiential content (of both types) evolves into the system of transitivity: transitivity and mood are the two fundamental components of the meaning of the meaning-making resources of every natural language. But we also see here something that is significant from the point of view of learning in general: namely, that construing experience is inherently an interactive process - there can be no content without also a speech function. The mood system is the resource for constructing dialogue; and it is only when the experiential content is mapped into a dialogic form that the child's world begins to

take semiotic shape. Commonsense knowledge is not a purely experiential construction; on the contrary, it is built out of the impact between the experiential and, the interpersonal modes of meaning. Learning involves both thinking and doing.

2. In the course of this impact, something else takes place. At the beginning of the transition from protolanguage to mother tongue, the child's mathetic utterances are, as it were, annotations, or footnotes to experience - a commentary on what is going on at the time, or an account of happenings from the past. They are not yet statements: that is, the child does not address these utterances to anyone who is not, or was not, a party to the happenings in question. Children may simply say these things to themselves. But if they are directed to another person, that person must be someone who is sharing or has shared the experience with them. Adults are frequently surprised to discover this; mother says, after an outing with her little boy, "Tell Granny what we saw" - but the child cannot do so. He may turn back to mummy, and tell her the whole story; but if he turns to look at granny, he is tongue-tied: - How can I tell Granny? She wasn't there. At this stage, language is a construction of shared experience - it is not a surrogate for it; and it is only when the two dimensions of meaning come together, when transitivity and mood combine to form a clause, that children can construe experience as news, using language not just to say but to tell. And once they can tell, of course, they can also ask. Again, when we trace the origin of telling and asking, we are looking at the child's development from a language point of view. What is the relevance to a general learning theory? It is that "information", something that we usually take for granted (it is after all built in to the concept of teaching), is not an inborn capability. Telling is a capability that has to be constructed - constructed in the course of learning language. It is only when you have learnt to tell that you can share experiences symbolically, as information, with those who have not been present with you to share in the events themselves.

The last two paragraphs have concerned developments that take place long before children go to school; they lie at the foundations of our unconscious, commonsense knowledge. There are other aspects of language learning which stretch out over much more extended periods of time. Let me turn next to two examples of these. The first I shall call the "interpersonal gateway".

3. I have suggested that language, in its distinctively human, adult sense, is an interplay of action and reflection: of the interpersonal and the experiential "metafunctions", in the terms of systemic functional theory. In every human language, whenever we speak (or write) we are typically at once both construing some aspect of experience and enacting some interpersonal force - the second of these includes both expressing our own angle on the matter and engaging in some relationship with another person, or other people. Both these components of meaning are present in all discourse. They are installed there by the grammar; hence, the grammar also makes it possible to foreground one or other of the two. It seems to be the case that when children are taking a major step forward in language learning they typically do so in contexts which are strongly loaded

interpersonally. One example could be drawn from my last heading, learning to tell: this step is likely to be taken under pressure from the expressive domain, when a child needs to convey that something unpleasant has happened - he has hurt himself, perhaps, and is needing sympathy. Another example, from a little later on, is that of learning to construe conditions: logical-semantic relations such as those expressed in English by *if, unless, although*. These are learnt in the first place, as Clare Painter (1989) and Joy Phillips (1986) have observed, in the context of threats and warnings: the adult says things like "if you touch the iron you'll hurt yourself", or "unless you stop banging that pan I shall take it away from you" - and the children then address such remarks to themselves, or to a younger brother or sister if one is available. In these and numerous other such examples, the meanings they are learning to make are primarily experiential in nature, semantic configurations which are going to play a central part in constructing knowledge, both commonsense and educational knowledge (like conditions); but the child's way in to these meanings is through the interpersonal gateway. And this again has implications for a general model of learning: the greater the conceptual distance that has to be traversed, in some particular learning task, the more critical it may be to set the task in an interpersonal environment - some context with which the learner is likely to be positively and interactively engaged.

4. The next feature is one which extends throughout the entire process of language development: the movement towards abstraction - children's progress through the semantic territory of the general, the abstract and the metaphorical. This too is a development in the potential of the lexicogrammar, and we can observe it as we track how children construct their grammatical resources. When they first move into the mother tongue, children learn to generalize: that is, they make the leap from "proper" to "common" terms - from naming individuals to naming classes, classes of things (persons and objects), of processes (actions and events) and of properties. These phenomena are construed in the open-ended word classes of every language, prototypically the nouns and the verbs. Children have no problem in construing as general terms the concrete domains of their "outer" experience: they readily master cups and dogs and buses, big and red, falling and hitting and breaking; and soon afterwards they also learn to construe their own "inner" experience of hurting and liking and remembering and seeing and so on. What they cannot yet cope with at this stage are words with purely abstract referents: words like *real* and *habit* and *choice* and *manage* and *delay*. Since one needs abstract meanings when learning to read and write (the teacher will often refer to *words* and *sentences* and *complete sense* and *information* and the like), it is at the age when children typically come to master this kind of language - round about five - that we put them into school. But it is not the actual skills of reading and writing so much as the entry into educational forms of knowledge that will make this demand on their language abilities. The primary phase of education depends on the learner being able to understand the meaning of abstract discourse.

But there is still another semiotic hurdle remaining to be crossed: the move from the abstract to the metaphorical. And this typically requires another four or five years of development. It is usually not until the age of eight or nine that children begin to accommodate metaphor in their grammar; and it takes them two or three years to sort it out and domesticate it. Now, while the educational knowledge of the primary school depends on abstractness, the discipline-based knowledge of the secondary school depends on metaphor: the sort of discourse that I illustrated earlier in the extracts from geography and history. Both the humanities and the sciences rely extensively on metaphor in the grammar, though in rather different ways. The history text talks about *war* and *peace* and *benefits* and *influences* and *supporting* and *promoting* and *progress towards a manufacturing basis*: these are metaphoric manipulations of abstract or institutionalized entities, which the learner has to relate to each other and assign appropriate connotations of value. The geography text talks about *withdrawal of heat, expansion, contraction, condensation, humidity, drainage, frontal uplift* and the like: these are processes and properties (get cooler, expand, shrink, condense, humid, drain, push up from the front) but they have been nominalized - that is, transformed metaphorically into virtual objects, the component parts of a systematic technical taxonomy. It is only by the time of adolescence that children are fully at home with this metaphorical mode of construing experience: when they move over from the primary stage of education into the secondary.

Thus it is the development of grammar that reveals most clearly the maturational principles that lie behind the structure of education - not only of educational knowledge but of the institution of education itself, the division of schooling into primary and secondary, with (in some systems) a middle or junior high school dedicated to helping children make the transition. Of course, the linguistic factors that I have picked out here as being critical in this developmental process are not suddenly appearing in isolation from everything else; they are part of the grammar's overall construction of experiential meaning. The grammar opens up a multidimensional semantic space through clusters, or syndromes, of related systemic features. To give just one example, at the same time as children are mastering these metaphorical nominalizations they are also, in English, developing the use of non-finite clauses, which are another element within the same area of semantic potential. But we can often identify certain specific components within the grammar which turn out to be critical for a particular "moment" in children's construction of knowledge.

I would like now to refer to three further linguistic features which illustrate my general thesis; but I will deal with them very much more briefly. They are, as those already discussed, aspects of children's language development which seem to me to offer pointers to the nature of learning in general. The three headings - somewhat opaque in themselves, but to be clarified, I hope, in what follows - are: the movement between system and instance; semiotic regression and reconstruction; and the synoptic / dynamic complementarity.

5. In learning language, children are all the time moving between the system and the instance. That is to say, they are construing the system - the potential of language, its semantic, lexicogrammatical and phonological resources - out of instances that they listen to and read; and, on the other hand, they are using these resources in speaking and in writing: from the system they are producing instances

of their own. It is the dialectic between these two that constitutes learning. We can often observe this movement when a child says something new, describing an event, perhaps, with a grammatical pattern that is extending the frontiers of his system; the child may then repeat the same account, many times, over the next few weeks and months, using precisely the same sounds and the same wording - by which time the system has moved ahead, and the instance now sounds like a fossilized relic of an earlier stage (which is exactly what it is).

6. In this particular case, there is no actual regression: the child's progress only appears to be stilted because we are hearing, at one and the same time, instances that were first worded at rather different times. But in one type of context there is a pattern of regression and reconstruction; this happens in the transition from commonsense to educational learning - it is an aspect of the discontinuity that I referred to right at the beginning. When children move into school, they face a considerable task of semiotic reconstruction: they have to reorganize their ways of meaning along new and unfamiliar lines. They have to re-form their language into a new medium, that of writing; and at the same time, or shortly afterwards, they have to restructure the discourse semantics so as to construe their knowledge systematically in a conscious form. In this process they often regress to earlier modes of meaning, on the one hand in their writing, so that a six-year-old, fluent and sophisticated in speech, will often write using the language of a child of two or three; and on the other hand in their understanding, so that they are learning over again things they already know perfectly well, but learning them now within an organized structure of knowledge. Children sometimes do not realize that something that is being presented to them in the written mode, and with all the majestic authority of the textbook, is actually something that has been part of their unconscious knowledge for some considerable time. I often cite the example from an upper primary science textbook, "Some animals protect themselves with bites and stings": in Australia, at least, children have known this since the age of two - it is important for their survival! They would not, of course, construe it in this way grammatically; they would say *they bite and they sting*, using verbs to express the actions, whereas the textbook is introducing them to scientific discourse and transforms these processes metaphorically into nouns: *bites and stings*. The experience is being reconstrued for them, by the grammar, as part of a different universe of knowledge.

7. And this leads me to the final heading, which I expressed technically (using grammatical metaphor) as "synoptic / dynamic complementarity". Here in fact this very fundamental notion of grammatical metaphor becomes central to the interpretation of learning. When children first construct the grammar of their mother tongue, they are able to do so very quickly because it provides them with a theory for explaining their own experience. So the structure of a clause, in English, or in Chinese, is a theory about actions and events; it provides (i) a class of words for the process that is taking place, the doing or happening - this we call a "verb"; and (ii) another, distinct class for the participants in the process, the persons and concrete objects that do things, or have things done to them - these are the "nouns". So the child construes a model of experience in which the basic unit is an action

or event, comprising a process and one or two participants, with the process represented as a verb and the participants as nouns. Thus the prototypical meaning associated with a noun is that of a person, animal or concrete object; that associated with a verb is doing or happening. Other aspects of the total phenomenon also have their typical forms of wording: adjectives construe properties, conjunctions construe logical-semantic relations and so on. Since the grammatical mode is clausal, which foregrounds doing and happening, the resulting picture of reality is a fairly dynamic one.

But later on, as we have seen, the grammar undergoes a change; it is reconstructed in different forms, with nouns, or rather nominal groups, taking over from clauses as the basis for organizing experience. Now if children's grammar had started out in this way there would be nothing metaphorical about it; the noun would have been the everyday, typical resource for talking about phenomena of every kind. But it did not. In their commonsense learning, nouns were names of things. The grammar is not now neutral any more; it is already semantically charged, and the nouns carry this semantic prosody with them wherever they go. So when experience is reconstrued, with educational discourse, into a nominalized form, this sets up a semantic tension, a complementarity of perspective. If students read about *evaporation*, and *seepage*, and *rainfall runoff*, in their hydrology text, these have the semantic features both of happenings, processes (water evaporates and seeps through, rain falls and then runs off) and of things, this being the prototypical meaning of a noun. We might want to say that no phenomenon can be both process and thing at the same time: the two are mutually contradictory. But that is precisely what *evaporation* and *seepage* and *rainfall runoff* are. Just one or two random instances by themselves would have no noticeable effect; but when the entire edifice of knowledge takes on this bivalent form it makes a profound difference to the learner's picture of the world.

The two conflicting forces, however, do not meet on entirely equal terms. Commonsense knowledge is deeply installed in our brains and in our bodies; but it is unrecognized - whereas the more lately developed perspective carries not only the full authority of educational discourse ("what the textbook says") but also the immense power of a knowledge that is organized and systematic: either in systems of values, typical of the humanities, or as in the sciences, where the grammatically constructed logical argument is further reinforced by the taxonomic resources of the lexicon. (Such taxonomies depend entirely on construing every phenomenon as a "thing".) The effect of this is to provide a less dynamic, more synoptic vision of the world, in which reality is as it were held still, rendered fixed, bounded and determinate, so that it can be observed, measured and, if possible, explained.

This suggests that we, as educators, need to be aware of the technical language of the scientific disciplines and to see it not as a "jargon", a set of unnecessary and often complex and cumbersome terms, but as a powerful grammatical resource with which experimental science reinterprets human experience. We might note here that technical taxonomies are rather less forbidding in Chinese than in English - whereas in its technical grammar, on the other hand, Chinese is the more problematic of the two (see Halliday & Martin 1993: chapter 7). But the implications for learning theory go rather further than this. It is not simply that we should be aware of how reality is construed in language, first in the language of the home and then later reconstrued in the languages of education. More especially, to a significant extent the process of learning consists in adopting complementary perspectives on experience: on seeing reality in ways which are at one level mutually exclusive, and even contradictory, and yet which taken together provide a deeper insight than either perspective adopted by itself. In one sense, the entire division into commonsense knowledge and educational knowledge, of which we tend to emphasize only the negative effects (and these there certainly are), may also have its positive function, if it is from the clash between these two very different modes of meaning that wisdom is ultimately attained.

I have made use of seven headings, as follows:

1. the functional multiplicity of grammar: "action & reflection"
    - enacting interpersonal relationships ["interpersonal"] and construing human experience ["experiential"]

2. "information" as dialogic exchange: "telling & asking"
    - combining mood (interpersonal) and transitivity (experiential) as the foundation of commonsense knowledge

3. the interpersonal "gateway" to learning
    - engaging with what is being learnt, through involvement of the "self" in interaction with others

4. the move towards the abstract
    - from generalization to abstractness to metaphor: creating new dimensions of semantic space

5. the dialectic of system and instance
    - construing grammar out of discourse, and construing discourse out of grammar

6. semiotic regression and reconstruction
    - accommodating the written medium, and reorganizing knowledge in systematic and conscious form

7. complementary perspectives on experience: "dynamic & synoptic"
    - maintaining the tension between reality as process (clausal) and reality as thing (nominal)

What these seem to suggest, if we put them together, is that learning, when seen from the vantage-point of language, is a highly complex endeavour - but one that is achieved through the interplay of a number of different meaning-making processes each of which by itself is rather simple. It is perhaps better to try and summarize them in a different order. (4) Children are progressively reconstruing experience, away from the immediate and concrete, using likeness (or analogy) to construe general categories, then abstract categories, then metaphorical categories. Adopting a topological framework we can say that each step creates, or rather allows the learner to create, new **dimensions** of semantic space. (7) The metaphoric categories require the learner to adopt simultaneously two complementary **perspectives** on experience. Three further factors also play a part in enabling children to learn: (2) knowledge first becomes dialogic, such that it is expanded by telling and by asking - the learner is **exchanging** meaning; (6) the learner often regresses and reconstructs, returning to the same experience at a "higher" semiotic level - the familiar phenomenon of **spiralling;** (3) major steps involve renewing connection with the self, and the axis of "you and me" - let us say that the learner is **engaging** with what is being learnt. (5) Throughout these processes the learner is always involved in the dialectic between the **system** and the **instance;** in language, this means building a grammar out of the discourse and building a discourse out of the grammar. (1) Finally, the concept of "language development" suggests that children are recapitulating, or re-enacting, the history of human knowledge - I do not mean modelling it in detail, but developing a semiotic, namely language, which is at one and the same time a mode of reflection and a mode of action. In other words, the learner is developing the **metafunctional** foundation on the basis of which knowledge itself is construed.

You may feel that considerations such as these are merely the abstract musings of a grammarian who (like the grammarian of folklore) is a dealer in symbols, far removed from the daily activities of the classroom. Some might think that nothing in the theory of grammar would be relevant to educational practice. But we are now educating the citizens of the twenty-first century; and the demands that are going to be made on their intellectual resources - their understanding of the world, and of their own situation within it - are truly formidable. The points I have raised here are my own perception of how aspects of the learning of language may relate to learning, and to teaching, in general. They may not be the main issues; they may even be wide of the mark. But if we want to understand how children learn, and how we, as teachers, can effectively contribute to this process, I think it can be helpful to explore a language development approach to education.

References

Bernstein, B. (1990). *The structuring of pedagogic discourse.* [Class, codes and control IV]. London and New York: Routledge.

Britton, J. (1970). *Language and learning.* London: Allen Lane, The Penguin Press.

Bruner, J. (1990). *Acts of meaning.* Cambridge, MA: Harvard University Press.

Halliday, M.A.K. (1993a). *Language in a changing world.* Canberra: Applied Linguistics Association of Australia [Occasional Paper 13].

Halliday, M.A.K. (1993b). Towards a language-based theory of learning. *Linguistics and Education* 5:93-116.

Halliday, M.A.K. & Martin, J.R. (1993). *Writing science: Literacy and discursive power.* London and Washington: Falmer Press. [Critical Perspectives on Literacy and Education].

Painter, C. (1989). Learning language: A functional view of language development. In P. Hasan, & J.R. Martin (eds.), *Language development: Learning language, learning culture.* [Meaning and choice in language, I]. Norwood, NJ: Ablex 18-65.

Phillips, J. (1986). The development of modality and hypothetical meaning: Nigel 1;72 - 2;72 University of Sydney Linguistics Department: *Working Papers in Linguistics* 3:3-20.

Wells, G. (1986). *The meaning makers: Children learning language and using language to learn.* Portsmouth, NH: Heinemann.

TEXT, TALK AND INQUIRY: SCHOOLING AS SEMIOTIC APPRENTICESHIP[1]

Gordon Wells

Introduction

The claim that language has a central role in all aspects of education, and particularly in learning and teaching, would not be treated as news by those whose job is to teach a first, second or foreign language. Furthermore, we should probably hope that the central role of language would be equally self-evident to teachers in all areas of the curriculum. However, if we were to try to specify in more detail what that role is or should be, we might rather quickly discover that there were some fundamental differences of opinion, corresponding at least in part to the 'subjects' taught, but also to what might be called teachers' underlying philosophies. Rather than spell out possible alternatives, however, I should like to present one particular conception of education in order to use it as a framework for considering the relationships between written texts, talk and action, as these apply in all areas of the curriculum.

Learning through Participation in Culturally Significant Activities

Many years ago, in *Education and Experience,* Dewey (1938) made a simple observation, but one whose implications for classroom practice have still to be taken seriously: Human beings learn by doing. And at about the same time, but in a very different part of the world, a similar insight about learning through joint activity was being developed into a comprehensive theory of learning-and-teaching by Vygotsky, his colleagues and students (Vygotsky 1978, 1987, Leontiev 1981, Luria 1978). Having remained unknown outside the Soviet Union until their work began to be translated into English and other languages in the last couple of decades, the key ideas developed by this 'troika' are now beginning to have a very considerable impact on education at all levels from pre-school to teacher education.

Simply put, Vygotsky argued that each human being's capacities for acting, thinking, feeling and communicating, although based in his or her biological inheritance, are crucially dependent on the appropriation of the practices and artifacts developed over time within particular cultures in the course of goal-oriented joint activity. As Budilova (1972:310) puts it:

> ...a special form of transmitting the achievements of preceding generations to the next takes place in human society; that is the achievements are embodied in the material and symbolic products of human activities, and specific human psychological abilities can be developed through the mastery of these products by each person.

Four features of this account merit further attention. First, the notion of 'product', or 'artifact'. The importance of material artifacts for the development of culture is by now well understood; the invention of the flint knife and, later, of the wheel are recognized to have radically changed the possibilities for action of the prehistoric societies which invented them and of those that took over their inventions. In more recent times, the same sort of significance is attributed to the invention of the printing press, powered flying machines and the microchip. But Vygotsky's great contribution was to recognize that an even greater effect resulted from the development of semiotic tools based on signs, of which the most powerful and versatile is speech. For not only does speech function as a tool that mediates social action, it also provides one of the chief means - in what Vygotsky (1987) called 'inner speech' - of mediating the individual mental activities of remembering, thinking and reasoning.

The second feature of importance is the strong emphasis on activity as the site of both invention and use of all forms of artifact. Whether material or symbolic, artifacts are embedded in practices which have as their object the satisfaction of perceived needs. In this sense, an artifact has no meaning out of the context of activity, and to master the use of an artifact is to learn to participate in the practices in which it plays a functional mediating role. Leontiev (1981:55-56) says in elaboration on this point:

> The tool mediates activity and thus connects humans not only with the world of objects but also with other people. Because of this, humans' activity **assimilates the experiences of humankind.** This means that humans' mental processes ... acquire a structure necessarily tied to the sociohistorically formed means and methods transmitted to them by others in the process of cooperative labor and social interaction ... In other words, higher psychological processes unique to humans can be acquired only through interaction with others, that is, through **interpsychological** processes that only later will begin to be carried out independently by the individual. (The emphases are as in original.)

In the last sentence of this quotation Leontiev focuses on the third and perhaps the most important feature of the sociocultural theory of intellectual development, namely that of appropriation. Simply put, all higher mental functions are dependent on semiotic artifacts and practices that are first encountered *intermentally* in purposeful joint activities, in which more expert members of the culture both demonstrate their use and assist the learner in mastering them. Through participation in which his or her performance is assisted, the learner gradually masters the practices in which these artifacts are used so that they also become a resource for *intramental* activity. As Vygotsky (1981:162) puts it, on its way to becoming an internal mode of activity, "any higher mental function necessarily goes through an external stage in its development because it is initially a social function."

However, appropriation is not the end of the process, for the final stage occurs in further action, when the learner makes use of the new function to participate more effectively in similar or related social activity. Appropriation of cultural artifacts and practices thus involves a contining three-stage cycle, to which corresponds a triple transformation. First there is the transformation of the learner - a modification of his or her own mental processes, that changes the ways in which he or she perceives, interprets and represents the world; second there is a transformation of the artifact itself, as its use is assimilated and reconstructed by the learner on the basis of the learner's existing knowledge; finally, in using the artifact to mediate further action, there is a transformation of the situation in which the learner acts which, to a greater or lesser degree, brings about change in the social practice and in the way in which the artifact is understood and used by other members of the culture.[2]

Finally, because appropriation of cultural resources takes place through the learner's participation in goal-oriented joint activities, a further key feature of learning concerns the part played by the other, more expert participants in facilitating this process. Their role is to help the learner to understand the significance of the activity as a whole and of the constituent actions and artifacts that mediate iits performance and, while taking responsibility for the organization of the overall structure, to involve him or her as fully as possible, providing help and guidance with those parts of the activity that he or she cannot yet manage on his or her own. However, this assistance is seen as only a temporary 'scaffolding' (Wood, Bruner and Ross 1976), the purpose of which is to enable the learner to become a fully competent, independently-functioning participant. Vygotsky (1978:90, 1987) described this 'teaching' role as working with the learner in his or her "zone of proximal development":

> an essential feature of learning is that it creates the zone of proximal development; that is learning awakens a variety of internal developmental processes that are able to operate only when the child is interacting with people in his environment and in cooperation with his peers.

To summarize this view of education, as it applies to schooling, therefore, we might characterize it as the creation of a collaborative community of practice, in which, through assisted participation in appropriate activities, students undertake a 'semiotic apprenticeship', as they individually reconstruct the resources of the culture as tools for creative and responsible social living in this and the wider community. In this characterization, it is activity which is central, for, to rephrase Dewey's insight, what we learn is what we do.

Learning in School: A Semiotic Apprenticeship

In choosing the metaphor of 'semiotic apprenticeship' to characterize learning in school, I wish to emphasize three important points. The first is that learning should be seen as the gradual but cumulative development of expertise through participation in the activities in which, in the various disciplines, knowledge is progresively constructed, applied and revised; and the second is that, in their learning, students should be assisted and guided by others who engage with them in these activities and share their expertise. Both these points have been developed in some detail by Vygotsky and by contemporary sociocultural theorists (Chang-Wells and Wells 1993, Cole 1985, Lave and Wenger 1991, Rogoff 1990, Tharp and Gallimore 1988, Wertsch 1985), so I shall not develop them further. However, my third point - the semiotic nature of learning - has received rather less attention, and so it is on this that I shall focus here.

In the preceding section, I frequently used the rather general expression 'the resources of the culture' to refer to what has to be learned in the course of the semiotic apprenticeship. Without attempting to specify in detail what these resources are, I should like to suggest that they can be thought of as consisting of three broad categories: a) attitudes and values concerning what are worthwhile activities to engage in; b) understanding of the practices involved in these activities; and c) mastery of the relevant artifacts and of the procedural and substantive knowledge associated with their use. In sociocultural theory, all these resources are viewed as mediational means for the achievement of collective and individual goals. In that sense, they can be thought of as 'tools'. But, whether material or symbolic, in order to perform their mediating function, all tools must meet two requirements: first, they must be capable of contributing to the achievement of desired effects in the world; and second, they must be in the hands of a person who understands their meaning and mode of functioning in relation to the goals of the activity they mediate (Cole, in press).

From the learner's point of view, the first requirement is not initially in question: the efficacy of a tool is taken for granted on the basis of its continued use in the home and wider community. In the first instance, therefore, the learner's major task is to discover - in action - when, where and how to use the culture's most important tools, that is to say, to learn their semiotic significance. In this process, language provides the most important resource of all, for it is, as Cole (in press) puts it, "the master tool" - the tool that mediates the learning of all others.

In order to attempt to explicate this central role of language in learning, I should like to draw on some of the ideas contained in recent papers by Halliday (in press, this volume), in which he summarizes and extends his previous work on language learning. In the introduction to this paper, Halliday states: "the distinctive characteristic of human learning is that it is a process of making meaning - a semiotic process; and the prototypical form of human semiotic is language." Then, in the body of the paper, he goes on to show in detail how, in learning the language of his or her community, that is to say in constructing its grammar as a 'meaning potential', the child appropriates a powerful and versatile tool for participating in and reflecting on activity, in collaboration with others. Like Dewey and Vygotsky, too, Halliday emphasizes that what children learn with respect to language depends on what they use it to do.

When we come to consider language from this latter point of view, it is clear that there are two rather different kinds of doing in which language plays a part, which we might gloss, rather generally, as 'acting' and 'understanding', corresponding to an emphasis on the interpersonal or the ideational metafunctions in terms of which - together with the textual metafunction - the grammar is organized (Halliday 1978). In the first of these, the primary function of language is to mediate action: to negotiate goals and means, to monitor other, non-verbal, forms of behaviour, and to manage the interpersonal relationships involved. In the second, where the emphasis is on reflection, the function of language which is emphasized is that of 'representing' objects, events and relationships - of creating "a theory of human experience" - "a semiotic world of its own: a parallel universe ... [which serves] as model, or metaphor, for the world of action and experience" (Halliday 1993b:15-16).

Of course, these two forms of doing are never totally distinct and separate. In the former, language provides an important means of making reference to the other components of the activity in which the participants are engaged; and in the latter, any sustained reflection also requires the management of goals and interpersonal relationships. This is true of discourse contributions of any size: Whether the utterance is a single clause or a full-length book, choices are made with respect to both ideational and interpersonal metafunctions; in all utterances, too, choices with respect to the textual metafunction are necessary in order to make the utterance effective in its context. Nevertheless, as Britton's (1970) rather similar distinction between the 'participant' and 'spectator' modes suggests, there is an important difference between the two types of doing with language - acting and reflecting - and, for each, the grammar as a whole provides rather different kinds of semiotic resources.

This distinction will become clearer, I hope, as we turn to the various functions that language performs in the different activities that we might expect students to engage in in the classroom, if these activities are seen as an apprenticeship into the various 'ways of knowing' (including knowing in action)on which the curriculum is based. For, in all the disciplines there are some activities that are more oriented to action and others that more obviously involve reflection, and each involves different choices from the semantic meaning potential.[3]

In the sciences, for example, a rather clear, although oversimple, distinction can be drawn between empirical and theoretical activities and, in each, language plays a rather different role. For example, we should expect to find that the kinds of discourse that occur in the context of planning and carrying out experiments are rather different from those that occur when predictions are considered in advance of the experiment or the significance of the results is subsequently interpreted in relation to the prevailing theory. In history, a comparable distinction can be made between the kinds of discourse involved in the obtaining and handling of the documents and other artifacts that constitute historical evidence, and those in which the significance of this evidence is evaluated and debated. Similarly, in the study of literature, there is clearly a difference between, on the one hand, reading a novel or poem and responding, as one reads, in terms of the particular thoughts and feelings evoked, and attempting, on the other hand, to explain the work's overall effect by reference to specific features of the text seen as instances of more general literary categories. In each of these cases, although the same substantive 'content' may be involved in the two situations, the activities in which this content is worked on are different, and so are the discourse genres through which these activities are enacted. In the first of the contrasted situations in each case, the discourse plays a somewhat ancillary role in relation to the activity as a whole, whereas in the second the discourse actually constitutes the activity.

The important point, however, is that, within each of these academic disciplines, both forms of semiotic activity are recognized to be important, as is mastery of the genres of discourse involved. Indeed, in achieving the overall goals of the discipline, the different discourse genres perform complementary and interdependent functions, together constituting, in large part, what it is to 'do' science, history or literature.

The same, I want to suggest, should be true in the classroom. In order for their learning to constitute a genuine apprenticeship into the different disciplines, students should have the opportunity to encounter and master the important genres of discourse in each discipline through engaging in as wide a range as possible of the activities in which those genres are used. They should also receive assistance in their learning in the form of appropriate models and constructive feedback and guidance.

Before going on to consider how such a goal might be achieved, however, I want to look more closely at the semiotic resources provided by language, starting with a consideration of some of the differences between talk and text.

The Distinguishing Features of Written Text

There are a number of ways in which spoken and written discourse can be contrasted, all of which are relevant to a consideration of the role of language in learning and teaching. However, because of limitations of space, I shall consider only three.

1. The abstract nature of written text

In the paper already referred to, Halliday focuses on the relatively greater abstractness of writing, with the concomitant demand this makes on learners to "reinterpret their experience in the new mode of written language" (1993b:109). Interestingly, in ***Thinking and Speech***, Vygotsky (1987:204) gives a very similar explanation of the intellectual development that is fostered by learning to read and write: "Written speech forces the child to act more intellectually ... It is a more difficult and more complex form of intentional and conscious speech activity."

As both these scholars point out, written language is more abstract than speech in three ways: first, it involves a 'second-order symbolism', with written symbols standing for the spoken 'words' of speech, which are themselves symbols; second, because the interlocutor is not physically present, the way in which meaning is communicated in writing is more abstract than in speech, since the message has to be realized through the lexicogrammar alone, unsupported by gesture and intonation, and without the opportunity to check understanding and supply additional information on request.

However, it is the third - the abstract nature of the meanings themselves - that is what often makes the written mode more difficult for children than speech. It is for this reason that Halliday describes the development of literacy as involving both the reorganization of the learner's grammar to handle the more abstract categories of written language and also the mastering of a new form of knowledge: "written, educational knowledge as against the spoken knowledge of common sense" (1993b:109).

2. The functions of written text

In order to understand why the meanings expressed in written language are frequently more abstract than those in speech we need to consider the different functions writing is used to perform. Although a written text can be read aloud and speech can be written down, the two modes are not interchangeable, nor are they typically used for the same purposes. In fact, the relationship between them is one of complementarity rather than of correspondence. To a large extent, this is because of the different media that they employ, and the manner in which each is produced and received. Most importantly, compared with speech, written texts have much greater permanence; they are also much slower to produce, and, in reception, much more under the control of the receiver. For all these reasons, extended written texts are particularly suited to activities involving individual reflection whereas, in many ways, the dialogic exchange of meaning characteristic of speech is more suited to collaborative action.

It is the relative permanence of written text that explains its earliest uses some 3000 years ago. For, from the evidence that is available, amongst the first functions that writing was developed to serve was that of an aide-mémoire with respect to important practical information to do with trade and taxes (Ong 1982). And, in a myriad forms, written text still performs this function in almost every sphere of contemporary life (Heap 1989). Then, from this beginning, we can trace the emergence of two further functions: the archival and the instructional. If records are kept, they can be gathered together and collated, giving rise to bodies of information that can be consulted on future occasions (Olson, in press). Similarly, procedural skills required in the performance of practical activities can be described in written texts so that the passing on of craft knowledge and other similar information is not dependent on transmission by example or word of mouth.

As can be seen, all these functions are very clearly tied to action. In a very obvious sense, the written text serves in each case as a tool, making information available when it is needed in the course of some kind of practical activity. It therefore follows that, to be competent in performing the activity, one needs to know when and how to use the tools of written text, as well as the other tools that play a part in the activity.

However, the uses of written text are not confined to what Heap (1989) refers to as the 'enabling' of practical activity. In every culture that has made widespread use of writing for practical purposes, two further, reflective, functions have eventually emerged. The first of these I have called the 're-creational', intending thereby to capture the dual function of literature as both the recreation in words of the experiences of humankind, both real and imaginary, and also the recreation of the writer or reader through engagement with such texts. Finally, there is what I have called the 'epistemic' function - the use of a text as a tool for thinking and developing new understanding, through the dialogue that takes place between the reader or writer and the text, as he or she struggles to construct meaning that is clear and coherent and, at the same time, consistent with all the available evidence, both in the text and in his or her experience (Wells 1990).

However, in distinguishing the different functions served by written texts, I do not wish to suggest that texts are strictly monofunctional. In principle, any of the above functions can be served by any text. Nevertheless, there is in practice - as with material tools, such as hammers or saws - a strong correspondence between text-types, or written genres, and the uses to which they are put. And these uses, in turn, are defined by the activities in which they occur.

To return to the apprenticeship metaphor, then, it is clear that induction into the different disciplines involves learning to use the written genres, both practical and reflective, which mediate the activities that constitute those disciplines. And, as the metaphor implies, this learning must take place, not as an independent, decontextualized event, but as an integral part of carrying out those activities.

3. Dynamic and Synoptic Representations

So far, in contrasting written with spoken discourse, I have emphasized the relative abstractness of the former. I have also drawn attention to the functions that are particularly well served by the permanence which is a defining characteristic of written texts, in particular the 'epistemic' function in which, through a dialogue with the written representation, the writer/reader can use the text as "a thinking device" for generating new meanings and refining those meanings that are already represented (Lotman 1984).

However, there is a third way in which written discourse frequently differs from spoken discourse, particularly in expository genres. And this is in the different ways in which experience is typically constructed in the two modes. Here is how Halliday (1993b:111) explains the distinction:

> A written text is itself a static object: it is language to be processed synoptically. Hence it projects a synoptic perspective on to reality: it tells us to view experience like a text, so to speak. In this way writing changed the analogy between language and other domains of experience; it foregrounded the synoptic aspect, reality as object, rather than the dynamic aspect, reality as process, as the spoken language does. This synoptic perspective is then built into the grammar of the written language, in the form of grammatical metaphor: processes and properties are construed as nouns, instead of as verbs and adjectives. Where the spoken language says **whenever an engine fails, because they can move very fast, ... happens if people smoke more,** the written language writes **in times of engine failure, rely on their great speed, ... is caused by increased smoking.**

The difference is not simply one of alternative modes of expression, however. As Halliday points out, corresponding to the grammatical differences are different perspectives on experience. In learning to read and write, therefore, children have to "learn to construe their experience in two complementary modes, the dynamic mode of the everyday commonsense grammar and the synoptic mode of the elaborated written grammar" (op. cit.:112).

However, Halliday is not the only one to have made this kind of distinction. Vygotsky (1987), too, in his account of the development of verbal thought, proposes a similar progression from 'spontaneous' to 'scientific' concepts. Spontaneous concepts, according to Vygotsky, are encountered and learned in the spoken discourse that occurs in the varied and naturally-occurring events of everyday living; the learning of scientific concepts, on the other hand, is dependent on schooling, and in particular on the use of genres of discourse - typically written - in which concepts are systematically related to each other through definition or explanation.

Yet another related distinction is that made by Bruner (1986) between the two modes of thought that he refers to as 'narrative' and 'paradigmatic'. The narrative mode is primary and, as he points out, it underlies children's early experience of conversation. It is a discourse of doings and happenings, of actions and intentions: agents act in the light of prevailing circumstances in order to achieve their goals. This is the dynamic perspective on experience to which Halliday refers and the basis on which Vygotsky's spontaneous concepts are constructed. It is a mode of discourse in which the grammatical organization of the clause corresponds to the 'natural' relationship between the entities, actions and circumstances in terms of which we typically describe and explain behaviour, our own and other people's. In fact, by analogy, even the smallest particles of inanimate matter are endowed with intentions and potential for action when viewed from this dynamic perspective.

However, the synoptic mode of written discourse - the 'paradigmatic', as Bruner calls it - has its value too. By recoding almost every aspect of experience - processes, attributes and relationships, and even complete events in all their detail - as nouns or nominal structures, it provides a way of symbolically managing the complexity and variability of experience, allowing it to be reconstrued in 'scientific' concepts, which can be systematically related in taxonomies; instances can then be counted, and made amenable to operations of mathematics and logic. Having its origins in the field of science and technology (Halliday 1988), this powerful discursive tool has, not surprisingly, been appropriated by other fields of inquiry and, in different forms, has come to play a major role in the written genres of exposition and argument in almost all the disciplines. Indeed, as Halliday (1993b:112) observes, "it invades almost every register of adult English that is typically written rather than spoken."

At the same time - as is being increasingly recognized - the nominalizing tendency of these 'genres of power' is proving a double-edged sword. While it may facilitate technological discourse, it does so by construing reality in a form that is remote from the dynamic perspective on experience that is embodied in everyday talk. And this makes these genres difficult for students to master. It also creates a division between the technocrats who control them and the uninitiated who do not - a division that, as Halliday (1993a:32) points out, "is certainly dysfunctional in a modern democratic society." There are good reasons, therefore, for the attempts that are being made in various domains to modify these written genres to bring them closer to the language of commonsense. However, for the foreseeable future, they are likely to remain the genres of power and, for that reason, students need to be given every assistance in appropriating them so that they can participate fully in the activities in which they are used.

The Complementary Roles of Talk and Text

In the previous section, the emphasis has been on how written text differs from spoken discourse. Now I want to consider how the differences between the two modes enable them to function in complementary ways in the performance of discipline-based activities.

If we review the contrasts that have already been discussed, together with a number that have only been implied, it can be seen that there is a considerable degree of parallelism between the relevant dimensions as demonstrated in Table 1.

Spoken discourse is - relatively speaking - more likely to occur in a social setting, in which several participants are engaged in dialogue about experience, often involving action - ongoing, or in the immediate past or future - which is viewed from a dynamic perspective. When people engage with written texts, on the other hand, they are more likely to be alone, reflecting on the matter addressed in the text through the medium of a genre which adopts a monologic and synoptic perspective on experience.

Table 1

The Complementary Roles of Talk and Text

Spoken	Written
Action	Reflection
Dynamic	Synoptic
Concrete	Abstract
Spontaneous	Scientific
Narrative	Paradigmatic
Social	Individual
Dialogue	Monologue

Adopting these stereotypical characterizations of the two modes, it is clear that, as tools, talk and text are best suited to mediate different tasks within any activity.[4] Talk is likely to mediate the planning, monitoring and evaluation of the actions to be performed, while published texts may be referred to in order to supply needed information for these tasks and written notes of various kinds made to record intentions and interim results; then, depending on the nature of the activity, it may conclude with the writing of some form of text that gives an account of what was done and why, and of what was achieved or learned. In this scenario, talk and text are complementary because they perform different functions in mediating tasks that occur at different stages in the activity.

What such a synoptic account fails to capture, however, is the more dynamic manner in which talk and text can complement and enrich each other through an exploitation of the intertextual relationships between them (Lemke 1992). For it is when participants move back and forth between text and talk, using each mode to contextualize the other, and both modes as tools to make sense of the activity in which they are engaged, that we see the most important form of complementarity between them. And it is here, in this interpenetration of talk, text and action in relation to particular activities, that, I want to suggest, students are best able to undertake what I have called the semiotic apprenticeship into the various ways of knowing.

In the next two sections I should like to explore the significance of such talk about texts from the perspective of two different, although related, kinds of learning that they facilitate. The first concerns the meanings that are being made - what Halliday calls the reconstruing of experience in the grammar of the genres of written text. In the second, it is the function of the text that is focused on - the part it plays in the larger activity, for it is in this context that issues of generic structure seem likely to be most readily understood. To illustrate the different opportunities for learning that talk about text can provide, I shall use a number of examples from work in science, recorded in one classroom over a period of two years.[5]

Building bridges between ways of knowing

First, let us consider the ways in which talk can help with the more abstract meanings that are found in the texts that students are given to read. For example, with texts that present new information or new ways of organizing relatively familiar information, it is possible to offer glosses - paraphrases, explanations or examples - that build bridges between the dynamic and synoptic perspectives and between the language in which each of these is expressed and the students' own experience. Here is a fairly typical example.

In the course of a unit on electricity, one of the nine-year-old members of the class has shown a persistent interest in electricity in the human body and has asked a number of questions about how, if electricity causes shock, it can also play a necessary part in the proper functioning of the body. To address this issue for Benjamin and the other children, the teacher reads a short passage from a book she has added to the classroom library and then provides a further gloss on the information in the text.[1]

T: I'm going to read this part, 'Electricity in the Human Body'because I know Benjamin is still not satisfied about it.
(reads) "Tiny electric signals, which can be called synapses, <travel> through the heart muscles, triggering and coordinating the heart beat. These signals send "echoes" (T adds: They put 'echoes' in inverted commas) through the body tissues to the skin. Here, they can be detected by metal sensors and displayed as a wavy line called the electro-cardiogram."

Now that is the most positive proof that the human body contains electricity .. Have you seen pictures - movies - where people are harnessed up to - and a person had a heart attack and you see this wavy line (demonstrating).

C: Yeah, yeah (excitedly)

T: Now those wavy lines are showing the electricity going through the human body . that's called the electro-cardiogram. And when a person is dead it goes 'deeeeeee'-

C: Yeah, a straight line

---

[1] In this and the following transcripts the following conventions apply: . = 1 sec. of pause; <> enclose segments were the transcription is uncertain; * = an inaudible word; CAPS = spoken with emphasis; underling = segments spoken simultaneously.

Here, the teacher has enabled the students to bring their own experience, whether first-hand or tv-mediated, to contextualize the less familiar language of the written text.

Earlier in the same lesson, the reverse strategy was used. With a chart showing the different sources of energy displayed on the board, the teacher asked the children to mention the sources of energy with which they were familiar in their homes and local community. Over the course of considerable discussion, their contributions were then built with the teacher's guidance into a more comprehensive and taxonomic account, in which a major distinction was made between renewable and non-renewable sources of energy. This co-constructed account was then used to make sense of the information presented in the text of the wall chart.

However, such discussions do not need to be restricted to occasions when the teacher interacts with the whole class. For example, in practical activities, where the students are working in groups and consulting a text that contains instructions for action, the meaning of the text is often clarified by matching the linguistic expressions with the materials available and with what the students already know about actions that may be performed on them. Similarly, in working with a historical text or a novel that a group is studying, discussion in which alternative individual interpretations are critically examined and compared can lead to a collaborative interpretation in the construction of which each participant extends his or her own personal understanding.

So far, I have considered talk about texts that are already written, where the problem is to recontextualize the meaning, often synoptically expressed, in the more familiar language of everyday speech. However, the same sort of facilitation can also occur when writers discuss the texts that they themselves are composing. For example, in negotiating what information to enter into a table showing the results of a practical activity, students are led to consider the structure of the table as a form of representation and, in the light of this, to decide what aspects of what they have done and observed it is appropriate to include.

The following example comes from the next stage of such an activity. Working in groups, students have tested a variety of materials to find out which can be used to complete a simple circuit, thereby making a bulb light up. Back in the whole class setting, the groups have reported their results, which have been written in a table in which a check mark has been entered either in the column for 'bulb lights up' or in that for 'bulb does not light.' Now the teacher asks them to draw some conclusions.

T: Now, just from this chart, what does it tell you?
Is something like paper a good conductor of electricity?
(Several children shake their heads)

So what are the materials that are poor conductors of electricity - that do not conduct electricity - where electricity cannot pass through there to get you a complete circuit?
What sorts of materials are not good conductors of electricity? ...
Can you name them? Marie?

M: Cotton reels

T: So what is it made of? Name the material .. did you observe? What's the cotton reel made of?

A: Plastic

T: Plastic . so plastic's not a good conductor, you see .
What other materials?

E: Um . beer bungs and . wood (almost inaudible)

T: Listen to the question . I want you to make a knowledge transformation, OK? I want you to transform what you did .
I asked what MATERIALS are not good conductors . I didn't <ask about the> beer bung

E: Wood

T: Wood, plastic . What else are not good conductors?

P: Paper

T: Now these are NON-CONDUCTORS (writing the word above the column in the table for 'bulb does not light up')
Now can you name for me the CONDUCTORS of electricity -the materials not the objects?

In this example, it is the table, as a genre of written text, that serves to focus the double reconstruing of experience, from the dynamic account of the specific objects tested to the more abstract representation in terms of materials, and then from the everyday categories of materials to the synoptic categories of conductors or non-conductors of electricity. And in this case, the teacher makes the process explicit by talking about the 'knowledge transformation' that the students need to make.

Later in the same unit on electricity, several instances of a somewhat similar process were observed among groups of students who were preparing posters to explain the functioning of various devices they had made, each involving some kind of electric circuit. Knowing that their posters would be on display at the science fair for other children in the school, they first discussed what they had done and why in everyday talk, then they decided how this should be expressed in a form suitable for display next to their working model.

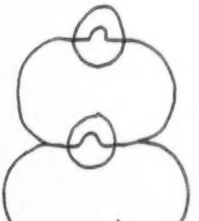
My final example involves a rather different kind of dialogue, but one introduced for the same reason - to encourage the students to use their texts as 'thinking devices' and, in the process, to build bridges between the two modes of representing their experience. From the beginning of the year, the students had been expected to keep a journal in which they wrote about what they had been doing and, more importantly, about what they thought they had learned in the process. In the electricity unit, a further dimension was added: two adults undertook to be journal correspondents, reading what the students had written and writing back with reactions and comments. The hope was that this would give the journal writing a communicative function that it had lacked before. And this indeed turned out to be the case. Many students addressed their entries to the adult partner and, in several cases, designated an area on the page for answers to their questions, including an injunction to the reader to respond.

Because of limitations of space, it is not possible to do justice to the range of genres and styles that were used. Nevertheless, the following example, involving a nine-year-old Chinese-Canadian student, gives an indication of the sort of exchanges that occurred and of the way in which the written journal dialogue helped the students build bridges between the dynamic and synoptic construals of experience.

March 23, 93. Dear Mr. Wells

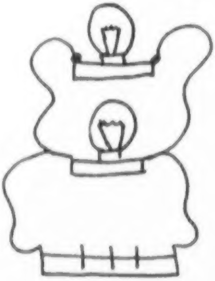
First today Mrs. Chang gave us an idea of drawing simbles. Then everybody gone to their groups and work on the light bulb. instead it had two bulbs and to make it difficul Mrs. Chang said if we can screw off one bulb off the bulb holder and see if the other bulb still lights up.

it was like this
with simbles



PARALLEL
CIRCUITS

it was like this
with out simbles

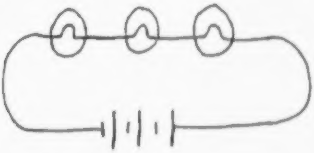


Dear Denny,

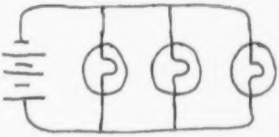
I liked the way you used symbols in your circuit diagrams.

I have a puzzle for you. Below, I have drawn two circuits. Each circuit uses a 6 volt battery and 2 volt bulbs. If you did this experiment, which bulbs would you predict would be most likely to burn out: (a), (b), both or neither?

(a)



(b)



I hope you will tell me what you think next time you write. Please explain the answer you give.

Gordon Wells.

Denny's Response:
The bulbs in (b) will be burn out. because the voltage drop on each bulb is 6 volts, but the bulb itself can only take 2 volts.

Making the form fit the function

In the examples of talk about text just considered, I drew attention to the ways in which collaborative discussion helps participants to clarify and develop their understanding of the 'content' of the texts they are working on and of the synoptic mode in which it is represented. However, as is clear from some of the examples, such discussion can also provide an occasion for learning about the equally important matter of 'form' and, in particular, about the structure of written genres.

As I argued earlier, an apprenticeship into the subject disciplines must include mastering the genres that are used by the members of those discipline-based communities in constructing and applying their 'theories of experience', for these genres are among the essential tools of their crafts.[6] Obviously, it takes many years to develop fluent control of the more specialized genres that are used in the publications of these professional communities, and such a degree of mastery is only achieved over the course of a career in which these genres are regularly used. However, as the notion of specialization suggests, there is a much smaller set of 'basic' genres, from which the more specific varieties are derived, and it is with these that the apprenticeship will naturally begin.

Considerable work has already been carried out in identifying key genres that provide an entry to the discourses of those disciplines that underpin the school curriculum (Martin 1993) and materials have been developed as a basis for systematic instruction in their use (Derewianka 1991, Christie et al. 1992). In Martin et al. (1987) genre is defined as "a staged, goal-oriented social process" and written genres are seen as particular instances of this more general category. In other words, written genres constitute the culturally developed ways of carrying out certain rather general communicative actions to achieve particular types of goals or purposes in the context of the relevant overarching activities. Thus, in terms of the sociocultural framework presented earlier, the different genres, such as 'recount', 'report', 'explanation', 'exposition' and so on, are best thought of as semiotic tools, whose use - as with other tools - is best learned when they are used to mediate the performance of the activity in question.

From this point of view, it may well be that materials that provide for systematic instruction in the use of particular genres can have a part to play in helping learners to understand the functional significance of their organization, particularly when the learning of the generic structure is embedded in a meaningful activity (Derewianka 1991). However, it is important to recognize that there are many other occasions on which such learning can occur in the course of the various activities that make up a curriculum unit. Indeed, in all the cases considered above where, in the course of carrying out an activity, the shared reading or writing of a text involved the co-construction of meaning, there was equally an occasion for discussion of the generic structure, as it related to the meanings made with the text and the purpose that the text served in the activity.

Here, there is space to include only one example, which occurred in the previous year in the same classroom as the previous examples. Jessica and Alan (both Chinese-Canadian children) had been conducting an experiment to verify that, as the text that they were using informed them, light is refracted when it passes through water. Pressed by a visitor to explain how their observations provided evidence of light bending, they had in fact gone beyond the experiment suggested in the book and devised a method of convincing both themselves and the visitor that the light beams, projected through two slits in the side of a shoebox containing a jamjar, only crossed when the jar was full of water. As Jessica triumphantly concluded at the end of the additional experiment, it was the water, not the glass, that made the light bend.

Later, the two children discussed the report they were expected to write with the visitor who had engaged in the experimental activity with them. First they considered who would read the report: people who already knew about refraction or those who didn't. Since their report was to be displayed in the classroom, they eventually decided that they would write for others who might be interested in carrying out the experiment. For this reason, they decided that they should start by 'writing the question'; then they should include two 'steps', one listing the materials needed and the other explaining what they did. There was considerable discussion about the ordering of these steps but Alan was finally persuaded by Jessica's argument:

> cos if you don't tell them the materials they sort of say if they want it- if THEY want to try it out um but don't know what the materials [are] they just- they sort of can't do it.

The remaining steps, it was agreed, would include what they observed and what they learned.

As these two nine-year-olds worked on their reports later in the day, there was further discussion between them about the actual layout of their report, as well as about the 'steps' they needed to include. Below, I have reproduced - as accurately as possible in the medium of print - the finished version of Jessica's report. What cannot be reproduced, unfortunately, is the series of annotated illustrations of the various stages of the experiment as it was actually carried out.

Bending light.

Question: Can light bend?

1 Materials: One cardbord box, a glass jar filled with water (to make exprimint more clearly put food coloring in water,) sheet of white paper that can fit inside the cardbord box, scissors, ruler, pen and a very bright flashlight.

2 What we did: First we drawed two nawrow slots two cm. apart each other on one side of the cardbord box. Then we cut the slots, put the sheet inside the cardbord box make sure it fits just right. Then we put the glass of water inside the cardbord box, make sure the jar of water is right beside the slots. In a very, very, very dark room (place), shine the flashlight through the two slots.

3 Observations: You might find out that you can only see two slots on the other side of the cardbord box but it doesn't mean that you did it wrong, if you don't belive me, try taking the flashlight and tip the back of it up (slightly) and then tip it so that it is leveled again. Repet that again and again and you will see it cross together, if you don't see it that means your either tiping it to much or you did something wrong. Now, say if you wearn't prety sure if it's crossing together and you want to be realy sure that it's crossed, try this, use one hand to hold the flashlight and one of your finger to cover one of the slots and then lift your finger up, now look at the other side of the cardbord box where the light will apear and do it again (lift your finger up and down) and if you would notice that when you cover the right slot the left slot will disappar and when you cover the left slot the right slot dissappars, you might wonder why, Because when you cover the right slot the right slot should dissappar, not the left, so this shows you that you may not see it cross but it mabe is.

4 Other questions people ask: Mr. X [Teacher B] asked Alan (my partner) and me a question, "What do you think is causing the light bend?" I said it was the water but my partner said it was the glass so insded we did another exprimint what we did was take out the jar of water and put in another jar but this time without water. we did exacly the same thing, and we tested it with our finger again, but it didn't crossover together. So we new it was the water.

5 Coments: I must say I have to thank [my teacher] for giving me an apertunaty to do this exprimint and learning so much things and also I have to thank [Teacher B] for helping us do this exprimint, thank you both of you. Another coment from myself, the exprimint was neat.

(Jessica)

Some might argue that, because Jessica's text is still an idiosyncratic mixture of several genres, it shows the need for a much more directive form of teaching. However, I would disagree. What is important, in my view, is not whether the texts that are the outcome of such collaborative discussion conform to some abstract prescription of 'report' or 'explanation' - a result that could be produced by filling in spaces in a pro-forma document - but that they have the form they do because the writers have made conscious decisions to construct them in that way in order to achieve the purposes that they set for themselves. Only when this is the case, I would argue, is it possible for them to use reader feedback as a basis for further discussion about whether the text is successful and, if it is not entirely so, about what sorts of changes could be made to improve it. Although this route to mastery may be slower, the advantage is that, at each stage, the learner is in control of the tool and can develop and adapt it to meet her expanding goals as a writer.

Talk, text and activity

With the preceding examples, I have tried to give an idea of some of the different ways in which talk about texts provides the occasion for simultaneously learning the new mode of written language and also the "written, educational knowledge" which is encoded in written texts. In concluding this section, I should like to review the points I have made from the perspective of the apprenticeship metaphor that I introduced earlier.

In sociocultural theory, as I explained above, learning is seen very generally in terms of appropriation. That is to say, learning is the taking over and internalizing of cultural artifacts and practices in the course of engaging in joint activities, in which the functional significance of these artifacts and practices is modelled and the learner receives assistance in their use. Talk almost always plays a part in this process, as participants discuss what they are doing and why. In the case of the appropriation of symbolic artifacts and practices, however, talk is absolutely essential (Wells 1990), since the way in which texts perform their mediating function is not as evident as in the case of the artifacts and practices that are used in such traditional, material activities as weaving (Rogoff 1990) or tailoring (Lave and Wenger 1991).

This is now well understood in the field of emergent literacy, where the widespread occurrence of collaborative talk about books and other texts in the preschool years has been well documented, as have been the benefits to be gained from these practices by young literacy apprentices (Crain-Thoreson 1993, Heath 1983, Teale and Sulzby 1986, Wells 1986). In the early school years, similar practices are found in the Reading Recovery programme developed by Clay (Clay and Cazden 1990) and in the 'instructional conversations' that are at the heart of the programs for minority students pioneered by Tharp and Gallimore (1988).

What needs to be emphasized, though, is that it is not only in the early years that learners benefit from working together to make sense of the texts they are reading and writing. The teacher members of my graduate classes have also found that they understand the readings better if they have the opportunity to discuss them in small groups (Wells 1994b). And the same is undoubtedly true for young adolescents in school, as they grapple with the unfamiliar forms and meanings of the synoptic genres of the subject disciplines in which they are expected to reconstrue their experience. However, as I hope to show below, it is possible to organize almost any curriculum unit in such a way that it provides multiple opportunities for the joint activities in which this sort of collaborative learning can occur.

Before leaving the topic of the complementary relationship between talk and text, however, there is one further point that needs to be made. And that is that, despite the characterisitic differences between these two modes of discourse in both form and function, there are also many intermediate forms that combine some of the features of each. Perhaps most important, in the present context, is the extended turn in dialogue, in which the speaker develops a topic in a systematic way, whether in narrating an event, describing a situation or process, or in stating and justifying a point of view. The oral expositions of new material that figure in many lessons - and which for many teachers are the prototypical form of teaching behaviour - are particularly clear examples of such intermediate modes. Their value is that they provide models of 'talking science' (Lemke 1990) -or mathematics (Forman, in press) or literature (Chambers 1993) - which introduce many of the features of the more formal written genres employed in these subjects, but in contexts in which the formal language is interspersed with the "everyday language of common sense."

Such models are undoubtedly important. But, as the main form of assistance, they are certainly not sufficient. Craft apprentices do not develop the skills they need simply by observing the artifacts produced by master craftsmen or even from watching the craftsmen at work. Certainly, the role of the master includes that of modelling the activity and explaining the principles and practices involved. But these contributions are of greatest value, not in the form of abstract precepts, but when offered as guidance and assistance as the apprentice is actually engaged in performing the activity (Collins et al. 1989). Similarly, in order to develop and hone their skills, semiotic apprentices also need guidance and assistance. But for them, too, this help is of greatest value when it is offered while they are at work on challenging projects that make constant demands on them to master the use of further tools and practices and even to invent new ones of their own.

It is important, therefore, that extended turns should not be the sole prerogative of the teacher. Indeed, as Lemke (1990) emphasizes, if students are to learn the genres in which scientists talk and write about the phenomena of interest to them, they need opportunities to do more than listen to teacher expositions or read what the textbook writers have written. They also need opportunities to talk and write science themselves, to others who are interested in, and responsive to, their contributions.

The same is equally true of other subjects, as is argued by Lampert (1992:307) about teaching mathematics to ten and eleven-year-olds:

> This means we do not proceed as if whatever the teacher says, or whatever is in the book, is what is assumed to be true. It also means that lessons must be structured to pursue the mathematical questions that have meaning for students in the context of the problems they are trying to solve. And this means that lessons are more like messy conversations than like synoptic presentations of conclusions.

In fact, "messy conversations" seems a very good way of describing those instances of talk about texts which, because directly related to problems with which students are grappling, are most productive for learning how both talk and text are used to make meaning and develop understanding. And when these conversations occur in the context of activities which the students have made their own, we have come close to optimizing, in school, the conditions under which these tools can be mastered. In the final section of this paper, therefore, I want to consider one way in which these conditions might be created.

Inquiry and Education

In my own learning as a teacher, one of the ways in which sociocultural theory has most helped me is in offering a way of reconciling the opposition that is often perceived to exist between the two overriding goals of education. These are, on the one hand, to ensure that the young are socialized into the values, knowledge and practices of the culture so that they grow up to be responsible and productive citizens and, on the other, to nurture the originality and creativity of individual students so that each is enabled to fulfill his or her unique potential. As teachers, we often find that, whilst believing in the second goal, the pressure to fulfil the first is so overwhelming that, in practice, there is little time or opportunity left to attend to the second. It is in this professional impasse that I have found the sociocultural metaphor of apprenticeship to be particularly helpful.

Two features of this metaphor, in particular, are worth exploring further. The first is the object of an apprenticeship. Certainly, it includes the passing on of the knowledge and skills of the craft, with an emphasis on application. This is the outcome that is emphasized in much of the current debate about accountability: What is learned in school should enable students to function effectively in the social and economic world beyond; theoretical knowledge is of value to the extent that it has implications for action. And there is much to be said for the argument that knowledge should be for effective action rather than simply for show under examination conditions.

However, there is more to apprenticeship than reproducing the achievements of the past. For the ultimate object is that the apprentice should become an independent master craftsman, who creates new artifacts and adds to the cultural resources. In fact, all of the inventions that we now take for granted grew out of past experimentation with the resources then currently available, as they were put to novel uses or adapted to deal with new problems in need of solution. In other words, creativity and originality are as much the object of education as is the reproduction of the existing order. Indeed, in the light of the problems facing humankind, they may be of even greater importance.

The second feature that needs to be explored is the means by which these twin objects are achieved. And here, unfortunately, the actual practices of trade guilds in the past leave much to be desired for, by all accounts, the young apprentice's life was very often one of drudgery and exploitation. These are not necessary conditions, however, and were probably as counterproductive in the past as they would be in any school or classroom today. On the other hand, the emphasis on learning through engaging in purposeful activity is as valid today as it ever was. And so is the principle of teaching by proposing challenging goals to be achieved, and providing assistance in meeting them in a form that is appropriate to the learner's needs and with the intention of enabling him or her to appropriate the practices that are enacted jointly, along with the responsibility for learning to manage them on his or her own.

This understanding of the teacher's role in assisting learning was expressed by Vygotsky (1978:87), somewhat aphoristically, as "what a child can do with assistance today she will be able to do by herself tomorrow." The passage occurs in the exposition of his conceptualization of teaching-and-learning as working with the learner in her 'zone of proximal development', that is to say, in the zone between what she can do alone and the upper limit of what she can do with appropriate help. What this means, in practice, is: engaging with learners in activities to which they are committed, observing what they can already do unaided; then providing assistance and guidance that helps them to identify the nature of their problems and to find solutions that enable them to bring the activity to a satisfactory completion. It is in this guiding role that the teacher can most effectively pass on the artifacts and skills developed in the past, for it is under these conditions that their utility is most evident and their mode of functioning most readily understood and mastered.

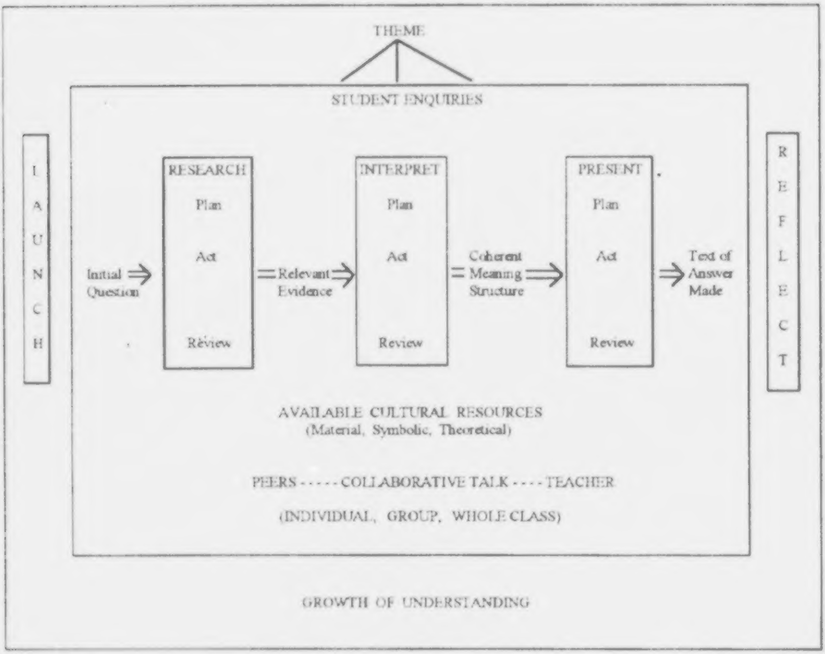
One way of organizing the curriculum to make this possible is by working with broad, open-ended thematic units, within which individuals or - even better - groups of students choose and plan their own topics of inquiry in consultation with the teacher. By selecting themes that both meet the requirements of the mandated curriculum and match the known or anticipated interests of the majority of the students, and by then sharing with them the responsibility for deciding on specific topics and how they should be investigated, the teacher maximizes the chances of achieving the first requirement - that the students should be engaged in challenging activities that they find personally significant. Under these conditions, student motivation is high and so is their ability to work independently, without the need for constant supervision and control. As a result, this mode of organization also meets the second requirement - that of allowing the teacher to be freed to spend time with individuals or groups, observing their progress and providing appropriate assistance when it is needed.

The first two requirements are concerned with ensuring that students' learning, and the provision of assistance, are embedded in a broader context of purposeful joint activity which, itself, involves a variety of constituent activities. However, there are two further, equally important requirements. The first is that, as well as being personally significant, the activities in which students engage should, over the unit as a whole, provide opportunities for them to make systematic progress towards mastery of the tools and practices of the discipline. This requirement can be met through the introduction of teacher-selected activities for the whole class, which are interspersed with the students' self-selected activities, and through the specification of genres to be included in students' presentations of their research. It can also be linked with the final requirement, which is that learning through action should be complemented by regular opportunities for learning through reflection. Whether undertaken individually or by the class as a whole, this will need to address both what has been achieved and discovered and the new questions that have arisen as a result, and also the means - the artifacts and practices - that have been employed in the process, as well as the problems encountered, whether solved or still in search of a solution. Whole class reflective discussion is particularly important here for, as well as fostering the development of the collaborative ethos of a community of inquiry, such discussion provides the setting, par excellence, in which knowledge is co-constructed, as students and teacher together make meaning on the basis of each other's experiences, supplemented by information from other sources beyond the classroom.

The overall structure of this inquiry-oriented approach to curriculum can be represented schematically as in figure 1. I must emphasize, however, that this is a 'tool' to be used for thinking and planning, not a prescription to be followed on every - or even any - particular occasion. Such decisions will always need to be made in the light of the curricular topic, the availability of resources of different kinds and, most importantly, the interests and capabilities of the particular class of students. (For further discussion of the model and its various components, see Appendix 1 and Wells (in press a).)[7]

Figure 1.

An Interactive Inquiry Model of Learning and Teaching



In the unit on electricity from which the earlier examples were taken, the *Launch* took the form of a brainstorming session in which each child first wrote down what he or she knew about electricity and their ideas were then discussed in a whole class session and written on a single large display, together with a list of questions to which students wanted to find answers (Scardamalia and Bereiter, in press). Then, following the viewing of a video-taped programme which provided an age-appropriate introduction to the topic of electricity, the teacher introduced the idea

BEST COPY AVAILABLE

of a science fair, for which groups would construct working models that involved some application of electrical circuits. The models that groups elected to make included a robot with flashing eyes, a truck-mounted electro-magnet and two morse code signalling stations connected by a length of wire.

While work proceeded on these artifacts, the teacher presented a series of challenges of increasing difficulty, starting with designing a simple circuit and continuing through parallel circuits with each bulb controlled by its own switch, to an experiment to discover the variables that determined the strength of an electro-magnet. Each of these activities was followed by a review session, in which the whole class *Interpreted* the results of the *Research* they had just done and *Reflected* on what they had learned, both about the principles of electricity and about strategies, social as well as practical and intellectual, for successfully solving problems of the kinds encountered. There were also whole class discussions, such as the one quoted from above, in which questions raised by the children were considered in relation to what was learned from the practical work as well as from consulting reference material of various kinds.[8]

By the end of the unit, then, the children had worked to find solutions to the problems encountered in constructing their chosen models and, through the teacher-posed challenges, had systematically learned about some of the basic principles of electrical circuits, conductivity and electro-magnetism. All this work culminated in a very successful science fair (*Presentation*), in which the groups' working models were complemented by displays of various kinds which, in every case, included posters providing explanations of the principles involved, as well as other information which group members thought would be of interest to the children from other classes who came to visit their fair.

As will be apparent, such an inquiry-oriented approach to curriculum creates opportunities for students to engage in many modes of discourse, both spoken and written. Earlier, I referred to several that I observed in the course of the curricular unit on electricity that I have just described. In Table 1, I have also drawn on my observations of thematic units based in different areas of the curriculum to present a more comprehensive account of the range of genres that might play a part as tools in mediating students' inquiries.

In presenting this summary, I have generalized across different patterns of participation. For example, in suggesting the genres of spoken discourse that might occur, I have not indicated the size of group that might be involved, as this might vary from a couple of students working together to the whole class. In the case of presentations, on the other hand, the audience might involve another class or parents and other adults invited for the occasion. Similarly, with respect to written texts: these might be produced by individual students or by collaborating groups and, where appropriate, addressed to a wider audience than the teacher alone through the use of class bulletin boards, or a class or school newspaper. In several cases that I have known, the final products of groups' inquiries, usually involving a variety of genres, were 'published' in book form and added to the resources in the school library. Finally, although I have only included 'journal entries' once, students might be encouraged to make regular entries in their journals or learning logs at all stages in their inquiries.

Looked at from the perspective of semiotic apprenticeship, Table 1 lists some of the more important discursive tools and practices that are utilized in carrying out an inquiry and indicates the tasks for which they are particularly useful. No doubt there are others that could or should be included, depending on the nature of the inquiry.

Whether spoken or written, however, the genres included in this table emphasize the inter-personal functions of discourse. Together, they make up a tool-kit for coordinating action and for negotiating and communicating participants' understanding with respect to their joint activities. However, they can also be seen as providing a similar resource for the intra-personal actions that participants carry on when they are alone, in what Vygotsky (1987) called the discourse of inner speech. From this perspective, participation in the genres of social discourse not only provides the means for the co-construction of knowledge; it also enables learners to appropriate the practices and artifacts and to use them to mediate the solo mental actions of thinking, imagining and reasoning. It goes without saying, of course, that these two broad functions - the inter-personal and the intra-personal - are interdependent and complementary, since they both mediate activities which, because they involve meaning, are inherently social and cultural. They are thus perhaps best seen as different phases of the continuing apprenticeship. As new genres of discourse are encountered and progressively mastered in inter-personal activity, they extend and transform the individual's intra-personal activity, and this, in turn, enables him or her to participate more fully and effectively in further inter-personal activity, in a never-ending spiral.[9]

Conclusion

In this paper, I have outlined a conception of education in terms of semiotic apprenticeship - an opportunity, through guided participation in discipline-based forms of inquiry, to appropriate the cultural tools and practices for meaning-making in the construction and application of knowledge in all areas of human activity. In this enterprise I have accorded a special place to language, seeing in the various genres of spoken and written discourse a kit of tools that performs a dual function, both mediating participation in activity and simultaneously providing a medium in which activity is represented and thus made able to be reflected upon.

In this conception - as befits the central metaphor - the emphasis is on the learner and on the conditions that enable him or her to master the means for full participation in the activity of inquiry, both alone and in collaboration with others.

Table 2

Genres of Discourse as Tools for Inquiry

	Oral Discourse	Written Discourse
Response to the Launch event	Drama, exploratory discussion, etc.	Poem, journal entry, etc.
	Brain-storming, formulating questions, hypothesizing, etc.	Statement of inquiry or of initial theory, etc.
Research	Planning, negotiating, coordinating and monitoring action;	Plan, list, instructions; letters requesting information;
	Observing, interviewing, consulting reference books;	Notes, tables of results, protocols, diagrams;
Interpretation	Interpreting evidence, debating alternative interpretations, drawing warranted conclusions;	Concept maps or webs, lists of arguments for and against alternative interpretations;
Presentation	Planning, negotiating, coordinating and monitoring action;	Plan, outline;
	Drama, report, video/audio program, commentary, panel, etc.	Narrative of events, procedural description, illustration, report, explanation;
Review	Reflecting, theorizing, evaluating;	Summary, reflection, evaluation.

As a consequence, it may appear that my intention has also been to deemphasize the importance of teaching. This is certainly not the case. However, an acceptance of this view of the classroom as a community of inquiry, in which learners share with the teacher the responsibility for deciding on the topics and on the means for their investigation, may indeed call for a reexamination of the ways in which the teacher's role might best be enacted (Wells 1994b).

However, in suggesting that we need to reexamine our conception of teaching, I am not arguing for the supplanting of one set of practices by a different set, that is already fully worked out and waiting to be applied. On the contrary: Every school and every classroom presents its own set of opportunities and constraints, and there is no set of practices that is guaranteed of universal success. Models, such as the one that I presented above, are no more than tools to be adapted in use to fit the prevailing conditions; it is to be expected, therefore, that they will be transformed by those who use them. The version represented in the examples quoted in this paper was developed by Gen Ling Chang to meet the needs of a culturally and socioeconomically diverse class of nine and ten-year-olds in Toronto. In other places, or with other grade levels, a different version might be more appropriate.

In other words, what I am suggesting is that teaching, like learning, is an ongoing process of inquiry, in which the knowledge that is constructed about learners and learning, as these are encountered in particular situations, continuously transforms the teacher's way of understanding and acting in the classroom. However, the practices of inquiry are not learned in isolation, nor do the various genres of discourse that mediate those practices take on their full value outside a context of joint activity. Like students in the classroom, therefore, teachers need to be participants in communities of colleagues who use the tools of inquiry to learn the craft of teaching (Wells 1994a). Furthermore, it is when we are ourselves intentional learners and inquirers that we most effectively model the practices that we wish our students to learn. For if what we learn is what we do - to rephrase Dewey's maxim - then what we do depends on the practices that are avilable for us to appropriate from the communities in which we participate.

1. The preparation of this paper was supported, in part, by a grant from the Spencer Foundation to the Ontario Institute for Studies in Education for a project entitled 'Learning through Talk'; however, the views expressed are those of the author and not necessarily those of the Foundation. I should like to thank the members of the Grade 4 and 5 class whose talk and texts are quoted in this paper both for allowing me to work with them and for what I learned from our collaboration. I should particularly like to thank the teacher, Gen Ling Chang-Wells, for the many discussions we had about her reasons for organizing the unit in the way she did and for her insightful comments on my interpretations of what I observed. Finally, I must acknowledge my debt to Glenn Humphreys and Patrick Allen for their constructive criticisms and comments on an earlier draft of this paper.

2. This triple transformation, as it applies in the appropriation and use of language as tool for both social and individual activity, is further discussed in Wells (in press b). As Halliday (1978) points out, each instance of language in use transforms the situation in which it occurs and either confirms or modifies the participants' view of the world, as this is construed in terms of the cultural categories encoded in language; by the same token, the language code is itself gradually transformed over time by the novel uses that speakers and writers make of it (Halliday, 1993).

3. In systemic linguistics, variation according to the use that is being made of language is handled in two related ways. Register accounts for the relationship between the social situation and the linguistic choices that are, so to speak, 'at risk' in it. Characterizing the situation in terms of three categories of properties - field, what is going on; tenor, the roles and statuses of the participants; and mode, the part that language is playing in the situation - Halliday (1975:126) defines register as "the range of meaning potential that is activated by the semiotic properties of the situation." Genre, on the other hand, accounts for the sequential organization of the stages by means of which the event is linguistically accomplished. For example, in a service encounter in a shop, there are five obligatory elements or stages - sale request, sale compliance, sale, purchase, and purchase closure - which must occur in that order (Hasan 1985). The discourse that actually occurs on any particular occasion can thus be accounted for by a combination of these two ways of thinking about variation (see Halliday & Hasan 1985, for a much more detailed exposition). In applying these two concepts to the context of education, Martin (1993) argues for the practical utility of giving primacy to genre, and this is the approach that I shall adopt here.

4. In Halliday's terms, all purposeful uses of language, whether spoken or written, involve the construction of text - discourse that is "functional in some context of situation" (1993b:107). This is an important point that should not be forgotten. However, in popular usage, the term 'text' is usually restricted to written discourse, which is distinguishable from spoken discourse in a number of ways, and not least - as I have suggested - by its relative permanence. Since this difference is particularly consequential for my argument, I shall stay with popular usage, referring to the two modes as 'talk' and 'text'.

5. This unit occurred in a Grade 4 and 5 class in an inner-city school in Toronto. Because of its location, the school serves a very diverse community; many of the children come from nearby Chinatown, but some are brought in from suburban homes

by parents who work in professional occupations in the adjacent hospitals and offices. The majority of children in this class spoke a language other than English at home and a few, being recent arrivals in Canada, were still in the early stages of learning the language of the classroom. As well as containing a number of children designated 'gifted', the class also included several children who were receiving help in part-time withdrawal programmes for literacy learning and behavioural difficulties.

As this study of electricity took place in March and April, the children had already had some experience of engaging in sustained inquiry, both in science and social studies. From the work on display in the classroom, it was clear that, wherever possible, connections were made between the topics chosen for inquiry and the children's activities in mathematics, literature, art and drama. The study of electricity should be seen, therefore as one theme within a curriculum that was both integrated and challenging.

6. While writing this paper, I came across the following paragraph, written by a professional scientist:

> Another problem with this ['scientific method'] paradigm is that it focuses only on the performance of experiments and overlooks that science is a social effort requiring communication. Because advances in science are interdependent, all the arts of communication are essential to science. Scientists visit one another's labs, travel to conferences, speak by telephone, hold advisory committee meetings, teach, argue, and write papers together, often using electronic mail. (Tinker 1993:2)

7. Not all applications of the apprenticeship model of learning provide so much scope for student initiative and creativity. Examples of more narrowly focused approaches are Brown and Palincsar's 'Reciprocal Teaching' (Palincsar and Brown 1984, Brown and Palincsar 1989) and Cole and Engestrom's (1993) 'Question-Asking-Reading'. An approach more like the inquiry model presented here is described in Gamberg et al. (1988), with many examples of themes that have proved successful with elementary age children. See also Kierstead (1985) for an overview of a similar approach to curriculum planning.

8. Benjamin, for example, searched among his father's books at home, consulted his uncle, a doctor, and accessed - with his teacher's help - the relevant entries in the CD-ROM Grollier Encyclopedia, in order to find out more about the role played by electricity in the human body.

9. There is one further feature of this table that only struck me after I had completed it: Quite unintentionally, I had for the most part chosen process terms to describe the genres of oral discourse, but product terms to describe the written genres. Whether or not this is appropriate, it does reinforce Halliday's point, quoted above, about the more synoptic, object-like nature of written text as compared to the dynamic, in-process nature of talk.

References

Britton, J. (1970). *Language and learning.* London: Allen Lane.

Brown, A.L. & Palincsar, A.S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L.B. Resnick (ed.) *Knowing, learning and instruction: Essays in honor of Robert Glaser.* Hillsdale, NJ: Erlbaum.

Bruner, J.S. (1986). *Actual minds, possible worlds.* Cambridge, MA: Harvard University Press.

Budilova, E.A. (1972). *Filosofskie probremy v sovetckoi psikxologii.* Moscow: Izdatel'stva (Nayka). [Philosophical problems of Soviet psychology.] Quoted, in translation, by K. Amano, E-mail XLCHC, 2 Jan. 1991.

Chambers, A. (1993). *Tell me: Children, reading and talk.* Stroud, Glos. UK: Thimble Press.

Chang-Wells, G.L. & Wells, G. (1993). Dynamics of discourse: Literacy and the construction of knowledge. In E. Forman, N. Minick & A. Stone (eds.), *Contexts for learning: Sociocultural dynamics in children's development.* New York: Oxford University Press.

Christie, F., Gray, B., Gray, P., Macken, M., Martin, J.R. & Rothery, J. (1992). *Exploring explanations,* Levels 1-4. Sydney: Harcourt Brace Jovanovich.

Clay, M.M. & Cazden, C.B. (1990). A Vygotskian interpretation of Reading Recovery. In L. Moll (ed.) *Vygotsky and education: Implications and applications of socio-historical psychology.* New York: Cambridge University Press.

Cole, M. (1985). The zone of proximal development: Where culture and cognition create each other. In J.V. Wertsch (ed.) *Culture, communication and cognition: Vygotskian perspectives.* New York: Oxford University Press.

Cole, M. (in press). A conception of culture for a communication theory of mind. In D. Vocate (ed.) *Intrapersonal communication: Different voices, different minds.* Hillsdale, NJ: Erlbaum.

Cole, M. & Engeström, Y. (1993). A cultural-historical approach to distributed cognition. In G. Salomon (ed.) *Distributed cognitions: Psychological and educational considerations.* Cambridge: Cambridge University Press.

Crain-Thoreson, C. (1993). The subtext of parent/child story reading. Western Washington University, Department of Psychology, unpublished paper.

Derewianka, B. (1991). *Exploring how texts work.* Rozelle, NSW: Primary English Teaching Association (Australia).

Dewey, J. (1938). *Education and experience.* New York: Collier Books.

Forman, E. (in press). *Forms of participation in classroom practice: Implications for learning mathematics.*

Gamberg, R., Kwak, W., Hutchings, M., Altheim, J & Edwards, G. (1988). *Learning and loving it: Theme studies in the classroom.* Toronto: OISE Press and Portsmouth, NH: Heinemann.

Halliday, M.A.K. (1978). *Language as social semiotic: The social interpretation of language and meaning.* London: Arnold.

Halliday, M.A.K. (1988). On the language of physical science. In M. Ghadessy (ed.) *Registers of written English: Situational factors and linguistic features.* London: Pinter (Open Linguistics Series), pp.162-178.

Halliday, M.A.K. (1993a). Language in a changing world. *Applied Linguistics Association of Australia,* Occasional Paper 13.

Halliday, M.A.K. (1993b). Towards a language-based theory of learning. *Linguistics and Education,* 5(2):93-116.

Halliday, M.A.K. (in press). A language development approach to education. In N. Bird. *Language and Learning,* Hong Kong: Government Printer.

Heap, J. (1989). Language as social action. *Theory into Practice,* 28:148-153.

Heath, S.B. (1983). *Ways with words.* Cambridge: Cambridge University Press.

Kierstead, J. (1985). Direct instruction and experiential approaches: Are they really mutually exclusive? *Educational Leadership,* May 1985:25-30.

Lampert, M. (1992). Practices and problems in teaching authentic mathematics. In F.K. Oser, A. Dick, & J-L Patry (eds.) *Effective and responsible teaching.* San Francisco: Jossey-Bass.

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* New York: Cambridge University Press.

Lemke, J.L. (1990). *Talking science: Language, learning and values.* Norwood, NJ: Ablex.

Lemke, J.L. (1992). Intertextuality and educational research. *Linguistics and Education,* 4(3/4):257-268.

Leontiev, A.N. (1981). The problem of activity in psychology. In J.V. Wertsch (ed.), *The concept of activity in Soviet Psychology.* Armonk, NY: Sharpe.

Lotman, Y.M. (1988). Text within a text. *Soviet Psychology,* 26(3):32-51.

Luria, A.R. (1978). *The making of mind: A personal account of Soviet psychology.* M. Cole & S. Cole. (eds.), Cambridge, MA: Harvard University Press.

Martin, J.R. (1993). Genre and literacy: Modelling context in educational linguistics. Sydney: University of Sydney, unpublished paper.

Martin, J.R., Christie, F. & Rothery, J. (1987). Social processes in education: A reply to Sawyer & Watson. In I. Reid (ed.) *The place of genre in learning: Current debates.* Geelong, Vic: Deakin University, Centre for Studies in Literary Education.

Olson, D.R. (in press). *The world on paper.* Cambridge: Cambridge University Press.

Ong, W. (1982). *Orality and literacy.* New York: Methuen.

Palincsar, A.S. & Brown, A.L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. *Cognition and Instruction,* 1:117-175.

Rogoff, B. (1990). *Apprenticeship in thinking.* New York: Oxford University Press.

Scardamalia, M. & Bereiter, C. (in press). Text-based and knowledge-based questioning by children. *Cognition and Instruction.*

Teale, W.H. & Sulzby, E. (eds.) (1986). *Emergent literacy: Writing and reading.* Norwood, NJ: Ablex.

Tharp, R. & Gallimore, R. (1988). *Rousing minds to life.* New York: Cambridge University Press.

Tinker, R.F. (1993). Curriculum development and the scientific method. *Hands On,* 16(2):2.

Vygotsky, L.S. (1978). *Mind in society.* Cambridge, MA: Harvard University Press.

Vygotsky, L.S. (1981). The genesis of higher mental functions. In J.V. Wertsch (ed.) *The concept of activity in Soviet psychology.* Armonk, NY: Sharpe.

Vygotsky, L.S. (1987). Thinking and speech. In R.W. Rieber & A.S. Carton (eds.), *The collected works of L.S. Vygotsky, Volume 1: Problems of general psychology,* (Trans. N. Minick). New York: Plenum.

Wells, G. (1986). *The meaning makers: Children learning language and using language to learn.* Portsmouth, NH: Heinemann.

Wells, G. (1990). Talk about text: Where literacy is learned and taught. *Curriculum Inquiry,* 20(4):369-405.

Wells, G. (1994a). Teacher research and educational change. In G. Wells et al., *Changing schools from within: Creating communities of inquiry.* Toronto: OISE Press, and Portsmouth, NH: Heinemann.

Wells, G. (1994b). Watching ourselves grow. In G. Wells et al., *Changing schools from within: Creating communities of inquiry.* Toronto: OISE Press, and Portsmouth, NH: Heinemann.

Wells, G. (in press a). Language and the inquiry-oriented curriculum. *Curriculum Inquiry.*

Wells, G. (in press b). The complementary contributions of Halliday and Vygotsky to a "language-based theory of learning." *Linguistics and Education.*

Wertsch, J.V. (1985). *Vygotsky and the social formation of mind.* Cambridge, MA: Harvard University Press.

Wood, D., Bruner, J.S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry,* 17(2):89-100.

CHINESE ORTHOGRAPHY AND READING: A CLARIFICATION

Ovid J. L. Tzeng

Writing systems have been qualified as logographic, syllabic or alphabetic according to the morphemic, syllabic or phonemic representation level of the speech (Hung and Tzeng 1981). Among the many writing systems existing in the world today, Chinese logographs are unique in that their relationship with the spoken language they transcribe is rather opaque. This relationship can be described as morphosyllabic in nature. However, the logographs and syllables do not have a one-to-one correspondence: the same syllable may be represented by different logographs with different meanings. The number of Chinese logographs has expanded to tens of thousands, and they are complex in configuration (Hung and Tzeng 1981, Tzeng and Wang 1983, Wang 1981). In his latest book on the developments of various writing systems, DeFrancis (1989) concludes, after a critical evaluation of the functional usefulness of most scripts, that a fully developed writing system has to be speech based. Interestingly, under such a conceptualization, Chinese writing system is very much sound-based and accordingly, its reading comprehension depends on the success of recovering its morphosyllabic representation. As we will see later, indeed, experimental results of recent psycholinguistic and neurolinguistic studies on reading Chinese are very much consistent with such an analysis (Tzeng, Hung and Lee 1991).

There is another unique aspect of Chinese logographs that needs to be mentioned. Centuries ago, these logographs were adopted by the Korean, the Japanese, and the Vietnamese to become their respective national writing systems. The sound systems of these languages are quite different from spoken Chinese, and there were major problems in adopting the Chinese writing system to transcribe them. Today, North Korea and Vietnam have dropped the use of Chinese logographs altogether and opted for an alphabetic system. However, South Korea and Japan have maintained them, and created sound-based systems (the Hangul alphabet for Korean and Kana syllabaries for Japanese) to overcome the problem of mismatch between the writing system and the sound system. Let us take a closer look at the Japanese case.

The origin of the Japanese spoken language is quite different from that of Chinese. The former evolved from the Altaic family of languages, which includes Turkish and Mongolian (Miller 1980). The latter, however, is not part of the Altaic group, and there are substantial differences in phonology between the languages. As a result of borrowing an orthography from a different spoken language, the Japanese have evolved two different pronunciations of the Kanji (the borrowed Chinese logographs) characters--a Japanese pronunciation and an approximation of the Chinese pronunciation. In addition, due to syntactic requirements, they have developed two syllable-based scripts in order to be able to represent function words and loan words. These are called Kana script in general, and the hiragana and katakana syllabaries specifically. Nowadays an ordinary Japanese text contains all three scripts in their distinctive styles.

68

For most Indo-European languages, the writing system, patterned after that of the Greeks, evolved to an alphabetic script, with the number of written symbols extensively reduced. A full alphabet, marking vowel as well as consonant phonemes, developed over a period of about 200 years during the first millennium B.C. in Greece (Kroeber 1948). The transition from the syllabic to the alphabetic system marked a gigantic jump with respect to the script/speech relationship. In fact, the development of vowel letters, which forms the basis of the analytical principle of an alphabetic system, has been characterized as something of an accident rather than a conscious insight (Gleitman and Rozin 1977). As a sound-writing script, an alphabetic system maps onto speech at the level of the phoneme, a linguistic unit smaller than the syllable but larger than an articulatory feature.

As we look back at these historical changes, we see that the evolution of writing seems to have taken a single direction: at every advance, the number of symbols in the script decreases, and as a direct consequence the abstractness of the relation between script and meaning increases and the link between graphemes and phonemes becomes clearer. This pattern of development seems to parallel the general trend of cognitive development in children and thus may have important implications for beginning readers of different orthographies. One of the major activities in learning to read is exploring the correspondence between the written script and the spoken language (Tzeng and Singer 1981). Since the script/speech relations in different orthographies tap into different levels of speech perception, and since the size of the minimal character set required for transcribing the entire speech segments in a language depends on such mapping relations, these unique historical developments provide ample opportunity to study the effects of orthographic variations on visual information processing within and across languages, and with respect to both skilled and beginning readers. A question of psychological interest concerns the extent to which different orthographies undergo similar (or different) processing. Only a correct description of the nature of symbols can help us to unravel the tangled story of success as well as failure in learning to read different scripts. Precise characterization of and closer examination into each type of the ever existing scripts in terms of the depth of script/speech mapping is necessary for any theoretical analysis of reading processes.

With respect to the question of linguistic relativity due to the variations in the orthographic structure, the Chinese language has been condemned as well as appraised, all because of its many unique properties. For example, in the 19th century August Schleicher proposed that "isolating" languages, such as Chinese, which used simple elements and were thus more "primitive" than" agglutinating" languages, which build their words from distinct forms. In contrast, as Wang has cogently pointed out, "Perhaps it is this structural simplicity of the language that moved the anthropologist and linguist Edward Sapir to characterize it as 'soberly logical'." (Wang 1973).

It is also true that more than any other writing system, the Chinese, with its non-alphabetic nature, has been besieged by "China experts" advancing potentially embarrassing notions. Most of these self-proclaimed experts are merely harmless drudges in the grip of a private theory. But there also were Leibniz and many other outstanding thinkers of the 17th and 18th centuries, who were much taken by the idea of creating a universal language based upon "scientific" principles similar to those which they thought underlay the Chinese system of writing (DeFrancis 1989). Such an idea persists even among modern-day scholars of high academic standing. Thus, for the well-known anthropologist Margaret Mead, the Arabic numeral system provides a partial model for a universal language of science, and the Chinese system "the most complete model" (Mead and Modley 1968:62). The enthusiasm was fueled by a research report in *Science* which showed that disabled readers of English in a Philadelphia elementary school were successfully taught to read English represented by Chinese characters (Rozin, Portsiky, and Sotsky 1971).

The results of the Philadelphia study and their implications have been disputed over the last two decades and the excitement of a possible "supreme orthography" dwindles down quite a bit after the observation of a null-finding from a renowned large scale study which involved three countries across three different writing systems (Stevenson 1984). However, curiosity about the on-line reading processes from a comparative perspective continues to persist among cognitive psychologists who are interested in building a "universal" theory of reading (Hung, Tzeng, Lee, and Chang 1994, Seidenberg 1985). The new debates center around topics such as the scriptal effects on the nature of reading disability, on modeling word recognition and naming processes, on the relationship between the phonemic awareness and learning to read, and on the development of higher cortical functions. A rigorous research methodology adopted from the experimental psychology tradition and a processing-oriented theorization imported from the emerging cognitive neuro-science program have helped to get rid of some of the wild notions about reading Chinese. In the following, I will review results from studies of the new approach under topics which are most relevant to our concern here.

Are Chinese Logographs Difficult to Learn?

One of the most popular statements made against the Chinese writing system is that, in spite of the beautiful shape and configuration of its logographic symbols, it is a very difficult writing system to learn for beginning readers. The belief in such a difficulty centers on the fear that the vast number of logographic symbols requires a tremendous amount of mental capacity for their mastery. A curious thing is that everyone seems to take this belief for granted and no one seems to be concerned with whether it is justified. There has never been a comparative study which examines the relative rates of reading acquisition in Chinese and another non-Chinese script (e.g. English). On what basis, then, can one make the claim that the Chinese is more difficult for learning to read? At a closer look, it can be shown that the specific negative statement made against the Chinese writing system is totally unfounded at both the theoretical and empirical levels.

Park and Arbuckle (1977) directly compared memory for words represented by Chinese logographs with memory for those same words represented in an alphabetic script (i.e. Korean Hangul letters). They found that the words were recognized and recalled more successfully when represented by logographs than when presented in alphabetic script. A similar conclusion was reached by Steinberg and Yamada (1979) in an experiment which compared the relative ease of learning Kanji and Kana symbols. They had 42 3- and 4-year-old children each learned a 4-item list for 10 trials with a paired-associate learning paradigm. The stimuli were 2 Kanji logographs and 2 Kana symbols and each of the four symbols were paired with its spoken name. That is, the subject's task was to learn to "recognize" the stimulus and supply its appropriate spoken name. The results showed that for the Kanji logographs, 37% were learned within 3 trials while most Kana were not learned at all even at the end of the 10th trial. Thus, young children seemed to learn Kanji faster than Kana script despite its perceptual complexity.

Of course, we cannot, and should not, take the above demonstration of the superiority of the logographic script too seriously either, for these studies are not without criticisms with respect to their methodology (see Tzeng and Singer 1978-79). Nevertheless, the point I want to emphasized here is this: No empirical evidence has ever been provided to support the proposition that the Chinese logographs are difficult to learn relative to other scripts! If anything, the results of those cross-script memory experiments, usually point to the direction favoring the Chinese logographs.

At a more extreme level, we can also take a serious look at the proportion of reading disabled children across different writing systems. In there, the issue of concern takes an ironical twist: In the literature on reading disability, the Chinese writing system has been appraised rather than condemned! That is, it has been suggested, with survey data as supporting evidence, that the occurrence of reading disability is extremely rare in the writing systems which use the Chinese logographs as their basic building blocks (Makita 1968, Kao 1978). Since the early 1980s, several groups of multi-disciplinary research teams, usually consisting of experimental psychologists, educators, speech pathologists, pediatric neurologists, and school psychologists, have carried out a series of comparative studies to test the validity of the commonly held belief that there are fewer incidences of reading disabilities in Chinese and Japanese.

The first serious attempt was launched in 1982 by Stevenson, Stigler, William, Lee, Hsu, and Kitamura. To achieve their research goal in three different languages (i.e. Chinese, Japanese, and English) and cultural environments, three compatible sets of individually administered and standardized Michigan Tests, which comprised graded reading tests and ten cognitive tasks, were constructed. They found that 6.3% children in Minneapolis (U.S.A), 7.5% in Taipei (Taiwan), and 5.4% in Sendai (Japan) could be classified as reading disability students. Thus, they concluded: (1) similar cognitive abilities were required to read all three languages, (2) visuospatial and perceptual abilities were not more involved in reading Chinese and Japanese than in reading English, and finally, (3) reading disability was as common in the two oriental countries as in the United States.

Taking advantage of the bi-literacy requirement in Singapore elementary schools, another group of investigators (Hung, Tzeng, Lee, and Chang 1994) examined the reading performance of Chinese-English bilingual children at the entry level. Their findings and conclusion were in general consistent with those of the Stevenson's group. Furthermore, Chang, Hung and Tzeng (1992) have done an error analysis of oral reading protocols of Chinese monolingual normal as well as disabled readers. A detailed analysis led them to the conclusion that most disabled readers of Chinese were not hampered by the learning of logographic symbols per se; rather, their problem lies in their inability to correctly segment the string of evenly spaced characters into appropriate word units. Why this should be so still remains an open question and future research has to pinpoint the exact nature of the word segmentation difficulty in reading Chinese from a theoretical perspective. Nevertheless, it is quite clear that once again we have to point out that there is no reason at all for the Chinese logographs to be casted as a difficult script to be learned.

Does Reading Chinese Require an Enhanced Visual Memory?

Studies in English have repeatedly demonstrated that reading ability is related to phonological memory and not to visuospatial memory (Baddeley 1986, Liberman, Mann, Shankweiler, and Werfelman 1982, Mann and Liberman 1984, Shankweiler and Crain 1986). Conventional wisdom says that in an logographic script such as Chinese characters (Hanji) the reverse should be expected because of the unique visuospatial arrangement in the printed symbols. But so far no direct evidence has been provided for or against such an expectation with respect to the reading processes of the Chinese writing system. The only study which has somewhat addressed this issue was conducted by Mann (1986). She tested 100 Japanese second graders on visuospatial memory for nonsense figures as well as phonological memory for Japanese nonsense words using a recurring recognition paradigm. She found that good readers performed better on both tasks compared to the poor readers and that there was a low but significant correlation between reading ability and phonological memory for nonsense words. In addition, the results showed that visuospatial memory for nonsense figures was significantly correlated with reading Kanji (adopted from Chinese Hanji) but not Kana (a sound-based script with each unit representing a syllable or mora).

Mann's (1986) data with the Japanese readers are intriguing, in particular the correlation between their visuospatial memory for nonsense figures and their reading performance with the Kanji characters. To explore these findings further, since 1988 two systematic studies have been conducted among monolingual and bilingual Chinese children to examine the relationship among reading ability/disability and linguistic and non-linguistic memory. One of the tasks for non-linguistic memory was adopted from Mann's (1986) study in which the nonsense figures were formed by abstract and nonsense patterns of lines and curves presented in a recurring visual recognition research paradigm (Mann 1986, Chang et al. 1992). Among the monolingual readers, the results showed that the disabled readers performed equally well as their normal achieving peers in both non-linguistic memory tasks, such as recognizing nonsense figures and reproducing geometric designs from memory. However, the disabled readers performed significantly lower than their peers in tasks requiring them to immediately recall ***in verbatim*** a sequence of digits and unrelated words. The literature shows that an inability to activate phonetic recoding, or representation, in order to maintain linguistic information in working memory is thought to be related to children with reading disabilities in English language (Brady and Shankweiler 1991, Leong 1991, Liberman et al. 1982, Shankweiler and Crain 1986, Torgesen 1988, Wagner and Torgesen 1987). Based on the comparative study between two groups of Chinese beginning readers, the results show that Chinese disabled readers performed significantly lower, when compared with their normal achieving peers, in processing language related elements such as digits and unrelated words (Chang et al. 1992).

Regression analysis of the scores obtained from bilingual children in Singapore suggests that phonological memory, measured through recurring auditory recognition of nonsense syllables, contributed towards prediction of English reading scores but not Chinese reading scores. Visuospatial memory, on the other hand, was not a significant factor in predicting reading in either language. Rote memory for shapes did not appear to be an important factor in reading Chinese logographs. Contrary to conventional wisdom, visuospatial memory ability could not explain the phenomenon of reading disability in both languages.

The results of a recent study in our laboratory tell a more revealing story. We were able to trace 16 students who were labeled "reading disabled" four years ago in the study conducted by Chang et al. (1992). These RD students had been assigned to a specially set-up resource classroom where they received intensive remedial programs in order to improve their reading skills. Among them, eleven had made satisfactory progress and were re-enrolled in the regular classroom. The remaining five, however, seemed to get nowhere even after the intensive training; they were probably all RD students in every sense of the word. Various experimental tests tapping on different cognitive skills were administered to these five RD children and to their peers with good reading skills. We found that on tasks that demanded the visuospatial memory ability, these RD children performed as well as their control group. They also showed no deficit in the memory for environmental sounds (i.e. wind blowing, car braking, whistling, etc.). Indeed, their cognitive deficit only showed up in speech-related tasks!

We shall discuss the issue of phonological processing in the next section. Here our concern is with the relationship between the visuospatial ability and learning to read Chinese logographs. Our conclusion is once again very clear: There is none! This is not to deny the importance of visuospatial processing during reading. Of course, it has to be an important component, because undoubtedly the visual sensory registration is the first stage of processing the printed symbols. However, there is no need to propose a special requirement for enhancing the visuospatial ability in order to learn to read the Chinese logographic system.

Does Skilled Reading of Chinese Not Require the Speech Recoding Process?

Fluent readers can read faster than they can talk, but for a child just learning to read, the opposite is usually true because every word has to be sounded out in order to get at the meaning. At some point during the process of acquiring reading skills, the transformation of visual code into speech code becomes automatic via some non-lexical symbol-sound correspondence rules, or becomes unnecessary altogether (the latter view has generally been referred to as the direct access hypothesis). In recent years, studies of word recognition in an alphabetic script like English have been dominated by concern over the nature of the code that allows the reader to go from print to meaning, a process called lexical access (Adams 1990).

Almost twenty years ago, when experimental psychologists started to launch their first series of attacks on reading from the perspective of information processing, using reaction time as the dependent measure, a number of investigators held the view that phonological recoding was a necessary preliminary to lexical access (Gough 1972, Gough and Cosky 1977, Rubenstein, Lewis, and Rubenstein 1971). A considerable amount of evidence was collected to support the phonological recoding hypothesis. However, other investigators were accumulating abundant evidence to support the direct access hypothesis. It is now clear from both the experimental and neuropsychological literature that, for a large number of words, phonological recoding for the purpose of lexical access is not necessary; in fact, some form of orthographic or visual code is sufficient for the purpose of getting meaning from print (Henderson 1982, Hung and Tzeng 1981, McCusker, Hillenger and Bias 1981, Saffran and Marin 1977, Seidenberg 1985).

Adding Chinese logographs into the picture seems to complicate, rather than clarify, the issue. Early supporters of the direct access hypothesis always used the example of reading Chinese to reinforce their argument. The argument goes like this: Because Chinese logographs do not contain information about pronunciation, people must be able to read without speech recoding.

However, this statement is not exactly correct. First, Chinese logographs consist of a majority of phonograms that at times do give clues to pronunciation. Thus, with the ability to pronounce a limited number of basic logographs, and knowledge of orthographic rules in the construction of logographs, readers of Chinese can in fact make reasonably successful guesses about how to pronounce logographs that share the same phonetic component, even those that they have never encountered before (Zhou 1978). The procedure involved in this type of grapheme-sound conversion is of course very different from that involved in the GPC (grapheme-phoneme conversion) rules advocated by Coltheart (1980). But it is similar to Glushko's (1979) activation-synthesis model of the generation of phonological codes. Indeed, such a procedure of generating phonological codes by analogy was proposed by Tzeng (1981) as one of two mechanisms in speech recoding, and was recently thought to be used by fluent readers of English for most words (Kay and Marcel 1981, Seidenberg 1985). Empirical evidence for the operation of this type of speech recoding in reading Chinese has been provided by Fang, Horng and Tzeng (1986) and by Lien (1985). Second, the Chinese writing system also makes it very clear that we cannot assume a one-to-one correspondence with respect to semantics between a word in print and a meaning in the mental lexicon. Single logographs are often recombined to make up new words; hence, there is nothing in the lexicon to be accessed. Meanings of words become available through the reference back to phonology and contexts. In this sense, it is rather difficult, if not impossible, to conceive of the access of lexicon via some orthographic or visual configurational cues. To a lesser degree this may also be true with respect to English orthography.

Reading should not be equated with the lexical access of a single word; rather, it should be regarded as a series of more general linguistic activities such as iconic scanning and storage, lexical retrieval, short-term retention, syntactic parsing at both macro- and micro-levels (Kintsch and Van Dijk 1978), and semantic integration over the entire discourse. This kind of conceptualization immediately questions the validity of the view that reading logographs involves no grapheme-phonology translation.

Thus, despite the bias towards direct grapheme-to-semantic processing, logographs may also activate phonological recoding processes. Erickson, Mattingly, and Turvey (1977) found increased errors in an immediate memory task when Kanji characters were phonologically related. Tzeng, Hung and Wang (1977) found similar effects in Chinese readers when phonetically similar logographs were used in an immediate memory task and in a sentence judgment task in which subjects decided whether sentences were meaningful and grammatically correct.

One implication to be drawn from all of these findings is that phonological recoding is just one of the strategies for obtaining access to meaning, rather than an obligatory stage. There are at least two major ways in which such a recoding process is important.

First, in blending the individual letters (or logographs) of words, the phonological recoding of the individual letter (or logograph) sound can plausibly be argued to be an important intervening stage, at least for children learning to read. A second way in which phonological recoding may be involved in reading is concerned with the question of whether fluent readers need to phonologically recode printed materials or are assisted by doing so. In this latter view the phonological recoding is regarded as a general strategy of human information processing, and thus, the orthographic difference in the printed materials becomes less important (Tzeng et al. 1977).

Is Phonological Awareness Not Essential for Learning to Read the Chinese Script?

Phonological awareness is the ability to recognize the internal structure of spoken words; it is usually assessed by testing the subjects' ability to isolate and manipulate individual phonemic segments in words. Much evidence is now available to suggest that awareness of the phonological constituents of words is an important prerequisite to fluent reading. This evidence comes from studies in several different alphabetic scripts which have shown that this awareness is predictive of reading success in young children (Adams 1990, Brady and Shankweiler 1991). No similar studies have yet been conducted in children learning to read logographic scripts.

Metalinguistic deficiencies in the phonological domain have also been demonstrated in adults with difficulty attaining literacy in alphabetic scripts (Morais, Carry, Alegria, and Bertelson 1979). However, a study in China found that adults literate only in traditional Chinese characters could not add or delete individual consonants in spoken Chinese words whereas adults literate in alphabetic Chinese as well as Chinese characters could (Read, Zhang, Nie, and Ding 1986). This study suggested that phonological skills involved in "segmentation" develop in the process of learning an alphabetic script, but not in learning a logographic script.

While there has been much evidence for the requirement of phonological recoding in fluent reading of Chinese (Tzeng et al. 1977), arguments against the idea that phonemic awareness may play a role in learning to read Chinese are still strong. This issue was examined in depth to compare the role of phonemic awareness in reading Chinese and English by studying two groups of Singapore beginning readers who simultaneously learned to read and write both Chinese and English (Lee et al. 1991). The findings of the first study showed that among the English-dominant bilingual children, their performance on a phonemic segmentation task correlated significantly with reading scores on both English and Chinese. It also was a significant predictor of reading ability in both languages. The findings of the second study among a group of children who were not dominant in English provided an interesting contrast. The relationship between reading achievement in English and phonemic awareness remained strong, whereas the relationship between reading Chinese and phonemic awareness became marginal. Together, these results suggest that it is alphabetic instruction, rather than maturation per se, that is responsible for the improvement in phonemic awareness occurring around the age children learn to read (Lee et al. 1991).

These two studies confirm the findings of other researchers who have studied children learning to read alphabetic scripts that phonemic awareness is important in reading English. However, the same issue is far more complex in the case of reading Chinese logographs. In our first study among the English dominant group, it is likely that these children gain the ability for analyzing the internal structure of speech sounds from learning to read English and in turn, use this ability to explore the phonological principles of Chinese logographs or characters.

It is well known that more than 85% of Chinese characters are phonograms. Each phonogram can be decomposed into two graphemic parts, a significate radical to indicate a general semantic category, and a phonetic component to give a clue to its pronunciation. Recent experiments by Tzeng and his associates have provided strong evidence that Chinese fluent adult readers take advantage of the generic properties of phonograms for decoding newly encountered Chinese characters. If this is the basic skill underlying the proficient reading of Chinese text, then it is likely that children who get access to this orthographic knowledge will be better able to expand their character size. Such a phenomenon has indeed been observed among Chinese monolingual beginning readers. Particularly, the speed in "character" acquisition among the disabled readers was impressive as they progressed through primary to intermediate grade levels (Chang et al. 1992). Hence, the unique formations of Chinese characters likely would not be the obstacle for reading and literacy development in Chinese logographic writing system. However, the Chinese "word" acquisition among the disabled readers lagged behind their normal achieving peers, as was evident in their reading error patterns.

In order to understand this line of research and the results obtained in the bilingual studies, the major issues are summarized as follows. First, the exploration of phonological clues from the Chinese characters is useful for reading. However, this presupposes that there is indeed phonological information available in the script, albeit some Chinese characters are more difficult to decipher. However, the connection between orthography and phonology is very important to all beginning readers. Early on Chinese children would have been exposed to some of the commonly used reading strategies to sound out unknown words. For example, if two graphic components are side by side, the strategy is to read the one on either side. If the character is formed by layers, the strategy is to try the sound clue presented in the center.

Second, the exploration of the script-speech, or orthography-phonology, relationship, though useful, is not the required way to learn to read Chinese because of the morphological differences. However, this is not to deny the importance of the role of phonological memory in the syntactic parsing and comprehension processes in which verbal elements are required to be held long enough to process information. On the contrary, since there is little pre-lexical phonological information available for the Chinese beginning readers to decipher logographs as opposed to sound-based alphabets, beginning readers have to rely solely on the post-lexical phonology, such as a learned pronunciation for each logograph or character, in order to convert the printed symbols into their phonological representation in memory. In a review of the literature, Adams (1990) concluded that such an automatic phonological processing ability is an important asset to all experienced readers. This may explain why the ability of phonological memory correlates with reading ability in Chinese.

Third and perhaps most importantly, the finding that the way a Chinese beginning reader acquires his/her reading skills can be influenced by the instructional environment may hold the key for the differentiation of alphabetic and non-alphabetic scripts. For students learning to read an alphabetic script, a pure graphic-based strategy, independent of phonology is not possible, whereas for students learning to read Chinese, which is morphosyllabic in nature, either the phonological or the orthographic strategy may predominate. Of course, as long as there is some phonological information embedded in the characters, there will be some overlap of these two options. It is suggested that for Chinese readers, the choice of either option depends a great deal upon the instructional environment, as revealed in two of these interrelated studies conducted in Singapore. Such a conceptualization may help to resolve much controversy on the necessity of "speech recoding" (e.g. converting the visual image of print into its phonological representation) in learning to read Chinese (cf. Tzeng and Hung 1988, Leong 1991).

Does Reading Chinese Involve a Greater Right Hemispheric Processing?

Throughout the history of hemispheric specialization research, there has been speculation about the possibility that the functional organization of a literate brain may be related to the type of written script one has learned to read. From Dejerine (1891) to Hinschelwood (1917), and from Luria (1970), Hecaen and Kremin (1976), Benson and Geschwind (1969), Zaidel and Peters (1981), evidence has been provided to show a selective sparing of reading one type of script despite severe impairments in the reading of other scripts in bilingual aphasic patients (for a more detailed review, see Hasuike, Tzeng and Hung 1986). Data from these bilingual studies are illuminating. However, they suffer from the lack of appropriate control of the degree of impairment of the spoken language. In this respect, recent findings of selective impairment in the reading of Kanji and Kana scripts by Japanese aphasic patients within a single spoken language have strengthened the hypothesis of the scriptal effect on cerebral organization (Hung and Tzeng 1981, Sasanuma 1980).

It should be noted that the finding of selective impairment in the reading of the two types of Japanese script does not necessarily implicate a right hemispheric involvement for processing Kanji. In fact, Sasanuma and her associates (Sasanuma 1975, 1980, Sasanuma and Fujimura 1971, Tatsumi, Itoh, Konno, Sasanuma and Fujisaki 1982) have argued for a differential disruption of language due to localized lesions in the left hemisphere, rather than postulating a dichotomy of right and left hemispheric processing for Kanji and Kana scripts. According to Hasuike et al. (1986), before the mid-1970s, there seemed to be no disagreement about the role of the left hemisphere for processing Chinese logographs. However, in 1977 two papers attracted much attention because both showed some evidence for right hemispheric involvement in reading Chinese logographs.

The first study was by Hatta (1977), whose results showed that native Japanese readers identified singly presented Kanji characters better when they were presented in the left visual field than in the right visual field, implying a stronger right hemispheric involvement. In previous studies (Hirata and Osaka 1967), native Japanese readers had showed the reverse lateralization pattern in identifying Kana symbols, implying a left hemispheric involvement in the processing of such sound-based script. Hatta's new finding was in accord with results obtained by Sasanuma, Itoh, Mori and Kobayashi (1977), in which nonsensical two-character Kana and Kanji characters were presented to native Japanese readers for identification. They found a significant right visual field superiority for the recognition of Kana symbols and a non-significant left visual field superiority for Kanji characters. Results from these two studies have often been cited to give evidence for right hemispheric involvement in the processing of Kanji logographs.

However, this seemingly clear picture begins to look very messy when one examines data from studies using Chinese readers. Visual hemifield experiments with Chinese subjects (Hardyck, Tzeng, and Wang 1977, 1978; Kershner and Jeng 1972) clearly showed a right visual field (left hemisphere) superiority for processing Chinese logographs. The discrepancy between the Japanese and Chinese results in these studies is curious. One possible interpretation is that Japanese readers process Kanji characters differently from the way Chinese readers process Chinese logographs, perhaps because of some unknown interaction between Kanji and Kana. Put another way, the Japanese not only borrowed the Chinese logographs, but also developed a different brain function in order to read them--hardly a plausible interpretation!

The major problem with visual hemifield experiments using a tachistoscopic procedure is the lack of control over the variables that could affect the results. Paradis, Hagiwara and Hildebrandt (1985) discuss such factors related to the nature of the stimulus, the presentation conditions, the task demands, the response, and the subjects, and note that in most studies the familiarity, concreteness, and types of logographs are often not specified, let alone be controlled. Thus, discrepancies could easily arise because of procedural differences. Tzeng, Hung, Cotton and Wang (1979) manipulated the number of logographs in two experiments, and found a left visual field superiority for recognition of single logographs and a right visual field superiority for two-logograph words. Hasuike et al. (1986) went a step further, in carrying out an extensive comparison among all relevant experiments up to 1985. They identified the stimulus exposure duration as the key variable because the left visual field's (right hemisphere) superiority was obtained only in those studies in which exposure duration was less than 50 msec. This makes sense: short exposure duration produces an incomplete visual image with a very low spatial resolution, and the literature has shown that the right hemisphere is adept in perceiving the relationship between these fragmentary components and the whole configuration (Sergent 1983). When the stimulus is presented for a longer exposure the spatial resolution is better, and under such conditions the left hemisphere seems to take over, especially when the task requires further linguistic analysis. It should be concluded then that there is very little evidence, from either experimental or clinical studies, to suggest a stronger right hemispheric involvement in the linguistic analysis of Chinese logographs. In fact, recent experimental evidence shows a very left hemispheric dominance in the processing of Chinese characters (see Bellugi, Tzeng, Klima and Fok 1990 for a critical examination of this issue).

Concluding Remarks

As we all know, Chinese is one of the written languages first recorded in human history. Fragmentary recordings date back more than 4,000 years, but real comprehensive recording came about only in the writings of the ancient sages, particularly the so-called confucian Classics, that appeared 3,000 to 2,500 years ago, followed by the even more extensive writings of the philosophers or latter-day sages, during the so-called Warring States period between 2,500 and 2,200 years ago. The confucian Classics, together with the writings of some of the philosophers, played a very important role in the formation of the Chinese civilization, perhaps a much greater role than that of the Bible for European civilization. It has been said that the Chinese civilization is like a crown shining through the history of humankind. It has further been said that the Chinese writing system is like a pearl on top of the crown. Interestingly but almost unbelievably, while the Egyptian hieroglyphs and Babylonian cuneiform can only be found in museums of history, modern day Chinese readers are reading exactly the same logographs as those read by their ancestors two thousand years ago.

In a sense, through the same medium, the Chinese people live simultaneously both in the present and in the past. It is no wonder that Chinese scholars are so proud of their writing system.

However, as we have shown in reviewing the current status of the Chinese writing system from the perspective of scientific research, widespread misconceptions about the nature of Chinese logographs have led to a lot of wild guesses about the psychological realities of reading Chinese. We have also witnessed an outpouring of incorrect claims made by famous neurologists and psychologists concerning the brain basis of reading Chinese. Of course, we should be delighted for the possibility that the Chinese language, due to its various unique features, has been considered to be an important language for a possible theoretical breakthrough in our understanding of reading behavior. But it is essential that every speculation about its psycholinguistic status has solid empirical foundations.

In this paper, we have carefully reviewed several issues pertinent to the learning and reading of written Chinese. Our conclusions are not much different from those reached a decade ago by Hung and Tzeng (1981), except that over the years, many minute details of the underlying operations have been identified and substantiated in experiments on a variety of cognitive tasks and with different subject populations (i.e. beginning as well as skilled readers, mono- as well as bilingual readers, RD students, and aphasic patients). It is clear that, despite the seemingly very different script/speech relationships embedded in different written languages, there is much commonality in the process of extracting meanings from print. The important question to be asked about reading, therefore, is not "what" are the differences?, but rather, we should ask why it should be that there remains so much commonality in the psycholinguistic processing across the perceptually very different scripts!

Acknowledgement

The writing of this paper was supported by a research grant from the National Science Council, ROC (NSC82-0301-H-194-018-Y).

References

Adams, M.J. (1990). *Beginning to read: Learning and thinking about print.* Cambridge, MA: MIT Press.

Baddeley, A.D. (1986). *Working Memory.* Oxford: Clarendon.

Benson, D.F. & Geschwind, N. (1969). The alexia. In P.U. Vinkel & G.W. Bruyn (eds.), *Handbook of Clinical Neurology, Vol. 1.* Amsterdam: North Holland.

Bellugi, U., Tzeng, O.J.L., Klima, E.S. & Fox, A. (1990). Dyslexia: Perspectives from sign and script. In A. Galaburda (ed.), *From neuron to reading: Toward a neurobiology of dyslexia* (pp. 137-171). Cambridge: Bradford Books/MIT Press.

Brady S.A., and Shankweiler, D.P. (eds.). (1991). *Phonological processes in literacy: A tribute to Isabelle Y. Liberman.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Chang, J.M., Hung, D.L. & Tzeng, O.J.L. (1992). Miscue analysis of Chinese children's reading behavior at the entry level. *Journal of Chinese Linguistics,* 20(1):120-159.

Chang, J.M., Lee, W.L., Loo, S. & Tzeng, O.J.L. (1990). A multiple-cue analysis on reading Chinese: Comparisons between elementary readers in Taipei and Singapore. Paper presented at The Third International Symposium on Teaching Chinese as Foreign Language, sponsored by Chinese Language Society, Beijing, China, August 15-20.

Chang, J.M., Rueda, R.S., Tzeng, O.J.L. & Bos, C.S. (1992). Locus of reading disabilities: Clues from a Chinese study. Unpublished manuscript. College of Education, San José State University.

Coltheart, M. (1980). Deep dyslexia: A review of the syndrome. In M. Coltheart, K. Patterson & J.C. Marshall (eds.), *Deep dyslexia.* London: Routledge & Kegan Paul.

DeFrancis, J. (1989). *Visible speech.* Hawaii: University of Hawaii Press.

Dejerine, J. (1891). Sur un cas de cecite verbale avec agraphie suivi d'autopsie. Mémoires Societé Biologique, 3:197-201.

Erickson, D., Mattingly, I.G. & Turvey, M.T. (1977). Phonetic activity and reading: An experiment with Kanji. *Language and Speech*, 20:384-403.

Fang, S.P., Horng, R.Y. & Tzeng, O.J.L. (1986). Consistency effects in the Chinese character and pseudo-character naming tasks. In H. Kao & R. Hoosain (eds.), *Linguistics, Psychology, and the Chinese language*. Hong Kong: University of Hong Kong Press.

Gleitman, L.R. & Rozin, P. (1977). The structure and acquisition of reading: Relation between orthographic and the structure of language. In A.S. Reber & D.L. Scarborough (eds.), *Toward a psychology of reading: The proceeding of the CUNY conference*. Hillsdale: Lawrence Erlbaum Associates.

Glushko, R. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 5:674-691.

Gough, P.B. (1972). One second of reading. In J.P. Kavanagh & I.G. Mattingly (eds.), *Language by ear and by eye*. Cambridge, MA: MIT Press.

Gough, P.B. & Cosky, M.J. (1977). One second of reading again. In N.J. Castellan, D.B. Pisoni & G.R. Potts (eds.), *Cognitive theory, Vol. 2*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Hardyck, C., Tzeng, O.J.L. & Wang, W.S-Y. (1977). Cerebral lateralization effects in visual half-field experiments. *Nature*, 269:705-707.

Hardyck, C., Tzeng, O.J.L. & Wang, W.S-Y. (1978). Cerebral lateralization of function and bilingual decision processes: Is thinking lateralized? Brain and *Language*, 5:56-71.

Hasuike, R., Tzeng, O.J.L. & Hung, D.L. (1986). Script effects and cerebral lateralization. In J. Vaid (ed.), *Language processing in bilinguals: Psycholinguistic and neuropsychological perspectives*. pp. 275-288, N.J.: Lawrence Erlbaum Associates.

Hatano, G., Kuhara, K. & Akiyama, M. (1981). Kanji helps readers of Japanese infer the meaning of unfamiliar words. *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition*, 3:30-33.

Hatta, T. (1977). Recognition of Japanese Kanji in the left and right visual field. *Neuropsychologia*, 15:685-688.

Hatta, T. (1981). Differential processing of kanji and kana stimuli in Japanese people: Some implications from Stroop-test results. *Neuropsychologia*, 19:87-93.

Hatta, T., Katoh, H. & Aitani, N. (1982). Does single kanji process dominantly in the right hemisphere? some implications from Stroop-test results. *International Journal of Neuroscience*, 18:67-72.

Hecaen, H. & Kremin, H. (1976). Neurolinguistic research on reading disorders resulting from left hemisphere lesions: Aphasia and "pure" alexia. In H. Whitaker & H.A. Whitaker (eds.), *Studies in Neurolinguistics, Vol. 2*. New York: Academic Press.

Henderson, L. (1982). *Orthography and word recognition in reading*. New York: Academic Press.

Hildebrandt, N. (1981). The Stroop effect with Japanese-English bilinguals. *Kinjo University Annuals, English Department*, 22.

Hinschelwood, J. (1917). *Congenital word-blindness*. London: Lewis.

Hirata, K. & Osaka, R. (1967). Tachistoscopic recognition of Japanese letter materials in left and right visual fields. *Psychologia*, 10:17-18.

Huang, C.Y. (1983). Varieties of deep dyslexia in Chinese orthography. Paper presented in the 6th Asian and Oceanian Congress of Neurology. Taipei, Taiwan, November 13-17.

Hung, D.L. & Tzeng, O.J.L. (1981). Orthographic variations and visual information processing. *Psychological Bulletin*, 90:377-414.

Hung, D.L., Tzeng, O.J.L., Lee, W.L. & Chang, J.M. (1994). Orthography, reading disability, and cerebral organization. In W.C. Watt (ed.), *Writing systems and cognition*. (pp. 11- 35). Amsterdam: Kluwer Academic Publishers.

Hung, D.L., Tzeng, O.J.L., Salzman, B. & Dreher, J. (1984). A critical evaluation of the horse-racing model of skilled reading. In H.S.R. Kao & R. Hoosain (eds.), *Psychological Studies of the Chinese Language*. Hong Kong: Chinese Language Society of Hong Kong.

Kay, J. & Marcel, A. (1981). One process, not two, in reading aloud: Lexical analogies do the work of non-lexical rules. *Quarterly Journal of Experimental Psychology*, 33A:397-413.

Kershner, J.R. & Jeng, A.G. (1972). Dual functional hemispheric asymmetry in visual perception: Effects of ocular dominance and post exposural processes. *Neuropsychologia*, 10:437-445.

Kintsch, W. & Van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review,* 85:363-416.

Klima, E.S. & Bellugi, U. (1979). *The signs of language.* Cambridge, MA: Harvard University Press.

Kroeber, A.L. (1948). *Anthropology: Race, language, culture, psychology prehistory* (new ed., rev.). New York: Harcourt, Brace.

Lee, W.L., Chang, J.M., Wee, C., Tzeng, O.J.L. & Hung, D.L. (November, 1991). Phonemic awareness and learning to read a logographic script. Paper presented at the 42nd annual conference of the Orton Dyslexia Society. Portland, Oregon.

Lee, W.L., Tzeng, O.J.L., Wee, G. & Hung, O.J.L. (1992). Intra- and Inter-language Stroop effects across three different scripts. In R. Harris (ed.), *Bilingual processing and cognition* (pp.427-442). Amsterdam: Elsevier Science Publishers.

Lee, W.L., Wee. C., Tzeng, O.J.L., and Cheng, L.O. (1989). The role of visuospatial & phonological memory in learning to reading Chinese. *Proceedings of the International Seminar on Chinese Language and Its Teaching in The World.* Singapore: Chinese Language Society.

Lee, W.L., Wee, G.C. & Wong, E. (November, 1987). Cognitive correlates of reading in English and Chinese amongst bilingual Chinese children in Singapore. Paper presented at the 38th annual conference of The Orton Dyslexia Society. San Francisco, California.

Leong, C.K. (1991). From phonemic awareness to phonological processing to language access in children developing reading proficiency. In D.J. Sawyer and B.J. Fox (eds.), *Phonological awareness in reading: The evolution of current perspectives.* (pp. 217-254). New York: Springer-Verlag.

Lien, Y.W. (1985). The reading process of Chinese character: The activation of phonetic information in the phonogram. Unpublished Master thesis, National Taiwan University.

Liberman, I.Y., Mann, V.A., Shankweiler, D. & Werfelman, M. (1982). Children's memory for recurring linguistic and non-linguistic materials in relation to reading ability. *Cortex,* 18:367-375.

Luria, A.R. (1970). *Traumatic Aphasia.* The Hague: Mouton.

Makita, K. (1968). The rarity of reading disability in Japanese children. *American Journal of Orthopsychiatry,* 38:599-614.

McCusker, L.X., Hillinger, M.L. & Bias, R.G. (1981). Phonological recoding and reading. *Psychological Bulletin*, 89:214-245.

Mann, V. (1986). Phonological awareness: The role of reading experience. *Cognition*, 24:65-92.

Mann, V.A. & Liberman, I.Y. (1984). Phonological awareness and verbal short-term memory: Can they presage early reading problems? *Journal of Learning Disabilities*, 17:592-599.

Marshall, J.C., Newcombe, F. (1973). Patterns of paraplexia: A psycholinguistic approach. *Journal of Psycholinguistic Research*, 2:175-199.

Mead, M. & Modley, R. (1968). Communication among all people, everywhere. *Natural History*, 77(7):55-63.

Miller, R.A. (1980). *Origins of the Japanese Language*. Seattle: University of Washington Press.

Morais, J., Cary, J., Alegria, J. & Bertelson, P. (1979). Does awareness of speech as a consequence of phonemes arise spontaneously? *Cognition*, 7:323-331.

Park, S. & Arbuckle, T.Y. (1977). Ideograms versus alphabets: Effect of script on memory in "biscriptural" Korean subjects. *Journal of Experimental Psychology: Human Learning and Memory*, 3:631-642.

Read, C., Zhang, Y.F., Nie, H.Y. & Ding, B.Q. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition*, 24:31-44.

Rozin, P., Poristsky, S. & Sotsky, R. (1971). American children with reading problems can easily learn to read English represented by Chinese characters. Science, 171, 1264-1267.

Rubenstein, H., Lewis, S.S. & Rubenstein, M.A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, 10:645-657.

Saffran, E.M. & Marin, O.S.M. (1977). Reading without phonology: Evidence from aphasia. *Quarterly Journal of Experimental Psychology*, 29:515-525.

Sasanuma, S. (1974). Impairment of written language in Japanese Aphasics: Kana vs Kanji processing. *Journal of Chinese linguistics*, 2:141-57.

Sasanuma, S. (1975). Kana and kanji processing in Japanese aphasics. *Brain and Language*, 2:369-383.

Sasanuma, S. (1980). Acquired dyslexia in Japanese: clinical features and underlying mechanisms. In M. Coltheart, K. Patterson & J.C. Marshall (eds.), *Deep Dyslexia*. London: Routledge & Kegan Paul.

Sasanuma, S. & Fujimura, O. (1971). Selective impairment of phonetic and nonphonetic transcription of words in Japanese aphasic patients: Kana vs. Kanji in visual recognition and writing. *Cortex*, 7:1-18.

Sasanuma, S., Itoh, M., Mori, K. & Kobayashi, Y. (1977). Tachistoscopic recognition of Kana and Kanji words. *Neuropsychologia*, 15:547-553.

Seidenberg, M.S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, 19:1-30.

Sergent, J. (1983). Role of the input in visual hemispheric asymmetries. *Psychological Bulletin*, 93:481-512.

Shankweiler, D. & Crain, S. (1986). Language mechanisms and reading disorders: A modular approach. *Cognition*, 24:139-168.

Singer, H. & Ruddell, R.B. (eds.) (1985). *Theoretical models and processes of reading* (3rd. ed.). Newark, DE: International Reading Association.

Smith, M.C. & Krisner, K. (1982). Language and orthography as irrelevant features in color-word and picture-word Stroop interference. *Quarterly Journal of Experimental Psychology*, 34A:153-170.

Stanovich, K.E. (1985). Cognitive processes and the reading problems of learning disabled children: Evaluating the assumption of specificity. In J.K. Torgesen B. Wong (eds.), *Psychological and educational perspectives on learning disabilities* (pp. 87-132). New York: Academic Press.

Stevenson, H.W. (1984). Orthography and reading disabilities. *Journal of learning disabilities*, 17(5):296-301.

Stevenson, H.W., Stigler, J.W., Lucker, G.W., Lee, S.Y., Hsu, C.C. & Kitamura, S. (1982). Reading disabilities: The case of Chinese, Japanese, and English. *Child Development*, 53:1164-1181.

Stroop, J.R. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology*, 18:643-662.

Tatsumi, I.F., Itoh, M., Konno, K., Sasanuma, S. & Fujisaku, H. (1982). Identification of speech, kana and kanji, and the span of short-term memory for auditorially and visually presented stimuli in aphasic patients. *Annual Bulletin of R.I.L.P.*, 16:205-218.

Torgesen, J.K. (1988). Studies of children with learning disabilities who perform poorly on memory span tasks. *Journal of Learning Disabilities,* 21(3):605-612.

Torgesen, J.K. (1989). Cognitive and behavioral characteristics of children with learning disabilities: Concluding comments. *Journal of Learning Disabilities,* 22(3):166-168.

Turvey, M.T. (1984). Investigations in a phonologically shallow orthography. Paper presented at the Third International Symposium on Psychological Aspects of the Chinese Language. Hong Kong.

Tzeng, O.J.L. (1981). Relevancy of experimental psychology to reading instruction. In O.J.L. Tzeng and H. Singer (eds.), *Perception of print: Reading research in experimental psychology.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Tzeng, O.J.L., Hung, D.L. (1981). Linguistic Determinism: A Written Language Perspective. In O.J.L. Tzeng & H. Singer (eds.), *Perception of print: Reading Research in Experimental psychology.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Tzeng, O.J.L. & Hung, D.L. (1988). Cerebral organization: Clues from scriptal effects on lateralization. In I-M. Liu, H-C. Chen & M.J. Chen (eds.), *Cognitive aspects of the Chinese language, Vol. I.* (pp. 119-139). Hong Kong: Asian Research Service.

Tzeng, O.J.L. & Hung, D.L. (in press). Location-specific inhibition in visual selection attention. *Psychological Science.*

Tzeng, O.J.L., Hung, D.L. & Lee, W.L. (1991, March). A conspiracy effect in the recognition of Chinese characters. Paper presented at the Conference on From Spoken to Written Languages, Bellagio, Italy.

Tzeng, O.J.L., Hung, D.L., Cotton, B. & Wang, W.S-Y. (1979). Visual lateralization effect in reading Chinese characters. *Nature,* 282:499-501.

Tzeng, O.J.L., Hung, D.L. & Wang, W.S-Y. (1977). Speech recording in reading Chinese characters. *Journal of Experimental Psychology: Human Learning and Memory,* 3:621-630.

Tzeng, O.J.L. & Singer, H. (1978-79). Failure of Steinberg and Yamada to demonstrate superiority of Kanji over Kana for initial reading instructions in Japan. *Reading Research Quarterly,* 4:661-667.

Tzeng, O.J.L. & Singer, H. (1981). *Perception of print: Reading research in experimental psychology.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Tzeng, O.J.L. & Wang, W.S-Y. (1983). The first two R's. *American Scientist*, 71:238-243.

Venezky, R.L. (1970). Regularity in reading and spelling. In H. Levin & J.P. Williams (eds.), *Basic studies on reading*. New York: Basic Books.

Wagner, R.K., and Torgesen, J.K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101(2):192-212.

Wang, W.S-Y. (1973). The Chinese language. *Scientific American*, 228(2):50-60.

Wang, W.S-Y. (1981). Language structure and optimal orthography. In O.J.L. Tzeng & H. Singer (eds.), *Perception of print: Reading research in experimental psychology*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Zaidel, E. & Peters, A.M. (1981). Phonological encoding and ideographic reading by disconnected right hemisphere: Two case studies. *Brain and Language*, 14:205-234.

Zhou, Y.G. (1978). To what degree are the "phonetics" of present-day Chinese characters still phonetic? *Zhongguo Yuwen*, 146:172-177.

TASK-CENTRED ASSESSMENT IN LANGUAGE LEARNING: THE PROMISE AND THE CHALLENGE

Geoff Brindley

1. Introduction

The advent of task-centred language teaching has brought with it various forms of assessment which are aimed at providing information on how well learners are able to mobilise language to achieve meaningful communicative goals. As Mendelsohn (1989) states:

> I believe that the goal of testing today ... is to see what someone **can do** with the language.

This type of 'can-do' assessment has a number of positive features:

- Teachers' and learners' attention becomes more focused on language as a tool for communication rather than on language knowledge as an end in itself (Shohamy 1992).
- Assessment is integrated into the learning process through the use of attainment targets which are directly linked to course content and objectives (Griffin and McKay 1992).
- Learners are able to obtain useful diagnostic feedback on their progress and achievement since explicit performance criteria are provided against which they can compare their performances. This fosters collaborative learning and encourages self-assessment (Brindley 1989).
- Better communication between users of assessment information and educational institutions can be established through the use of various forms of outcome reporting which are couched in performance terms and are hence intelligible to non-specialists (Griffin and Nix 1991).

However, despite these attractive features of task-centred assessment (TCA), it is not without its problems. Concerns have been have expressed regarding the validity of some forms of TCA (Bachman 1990), the feasibility of achieving reliability (Swain 1993), and the practical constraints, particularly the financial costs, involved in its implementation (Shohamy 1993). Some applied linguists have argued that a lot more research into the theoretical foundations of TCA is needed before this kind of assessment can be automatically endorsed as a viable alternative to more traditional forms of assessment (McNamara 1990).

In this paper I want to examine some of the issues and problems which have been raised in relation to TCA in the context of language learning. I will deal with these under the broad areas briefly mentioned above: *validity*, *reliability* and *practicality*. In the first part I shall examine some of the controversies surrounding the notion of authenticity as it relates to the validity of TCA. I will then discuss the important question of the construct validity of TCA by looking at various ways in which the criteria for judging task performance have been established and then suggest ways in which these criteria might be made more accountable to empirical data derived from studies of language acquisition and use. Since the success of TCA depends very much on expert human judgement, the second part of the paper will focus on some of problems which have been encountered in attempting to ensure reliability in TCA and look at the potential contribution of recent advances in measurement technology to addressing these problems. Finally, since TCA can only work 'on the ground' if the conditions exist for its implementation at a system level; in the final part of the paper I shall explore some of the practical ramifications of introducing TCA systems into educational institutions.

2. Defining Task-Centred Assessment

As various authors have pointed out, the term 'task' is used in a variety of ways in the language learning literature, ranging from very broad definitions that accentuate the 'real-world' nature of tasks:

> ... the hundred and one things people *do* in everyday life, at work, at play, and in between. Tasks are the things people will tell you to do if you ask them and they are not applied linguists (Long 1985:89).

to those that focus primarily on the role of language in the classroom:

> A range of work plans which have the overall purpose of facilitating language learning-from the simple and brief exercise type to more complex and lengthy activities such as group problem-solving or simulations and decision-making (Breen 1987:23).

For the purposes of this discussion I will propose the following definition of task-centred assessment which is sufficiently general to cover both in-class and out-of-class situations and can thus be applied to the assessment of proficiency acquired independently of the curriculum or to curriculum-based achievement:

> Task-centred language assessment is the process of evaluating, in relation to a set of explicitly stated criteria, the quality of the communicative performances elicited from learners as part of goal-directed, meaning-focused language use requiring the integration of skills and knowledge.

This definition draws on definitions of task-centred language learning as enunciated by, *inter alia*, Crookes (1986), Nunan (1989) and Swales (1990), and attempts to incorporate some key notions of criterion-referenced communicative assessment, namely:

- the need for explicitness in stating the criteria according to which learners' performances are to be judged (Brindley 1991).
- the centrality of communicative goals as a starting point in syllabus design and assessment (Nunan 1993).
- the view of language proficiency as encompassing both knowledge and ability for use (Bachman 1990).

TCA, of course, is not new in language learning. As far as proficiency assessment is concerned, the oral interview and various kinds of writing tasks have been a standard part of tests of communicative language proficiency for some time. And with the advent of the task-based syllabus, assessments and diagnostic feedback based on teachers' observations of task performance have become a feature of classroom practice (Brindley 1989). What is relatively new, however, is the weight which is now starting to be attached to assessment of learners' ongoing task performance as a factor in final assessment for purposes of certification. Another comparatively recent change in the assessment landscape is the degree of explicitness and rigour with which teachers are now being required to document and justify their assessments (Barrs 1992). What was once informal and formative, in other words, is becoming a high-stakes business. For this reason, TCA is being increasingly called to demonstrate its validity, reliability and practicality in much the same way as standardized pencil-and-paper tests have been in the past, and it is to the first of these issues that I now turn.

3. Validity Issues in Task-Centred Assessment

3.1 Validity and authenticity

One of the most attractive aspects of adopting the communicative task as a unit of teaching and assessment in language learning is that it enables the teacher to focus on communicative activities resembling the authentic use of language such as listening to news broadcasts and picking out the main points, reading a TV guide to find out what's on, defending a point of view, writing letters to a pen-friend in a foreign country, etc. These tasks can readily be turned into assessment activities as long as they are accompanied by a set of assessment criteria which describe what the learner must do in order to demonstrate that he or she is able to perform the task successfully.

It might seem reasonable enough to assume that assessments based on communicative tasks such as these are valid by definition since they attempt to replicate 'real life' language use situations, which is ultimately what communicative language teaching and assessment are concerned with. However such assumptions have been questioned by many writers in both general education and language assessment on a number of grounds. In the first place, an assessment activity is by its very nature an artificial situation: no matter how 'life-like' the task is, people still know they are being assessed under special conditions. As Spolsky (1985:36) comments:

> ... We are forced to the conclusion that testing is not authentic language behaviour, that examination questions are not real, however much like real-life questions they seem, and that an examinee needs to learn the special rules of examinations before he or she can take part in them successfully.

A second problem with 'authentic' assessment tasks is the difficulty of generalizing from a one-off performance to other situations of language use. Commenting on the tendency of some language testers to claim validity on the basis that their tests reflect real-life settings, Skehan (1984:208) comments:

> This viewpoint confuses naturalness of setting with sufficiency. A large part of the problem in testing is in sampling a sufficiently wide range of language to be able to generalize to new situations. Merely making an interaction 'authentic' does not guarantee that the sampling of language involved will be sufficient, or the basis for wide-ranging and powerful predictions of language behaviour in other situations.

In a similar vein, Bachman (1990) has cautioned against the acceptance of authentic-looking 'direct' tests such as the oral interview as automatically valid measures of ability. He points out (1990:309) that such tests confuse the observation of a performance with the ability itself and are limited in their generalizability beyond the specific context in which testing takes place. Bachman proposes a somewhat different approach to authenticity by suggesting that authenticity lies not only in the surface resemblance between assessment tasks and real-world behaviour but also in the extent to which different areas of language skills and knowledge are sampled in the task. This is what he refers to as ***interactional authenticity***:

> In summary the IA (interaction/ability) approach views authenticity as residing in the interaction between the test taker, the test task and the testing context. (Bachman 1990:317)

In order to construct valid 'authentic' tests of communicative language ability Bachman argues that we have to construct or select tests or test tasks that reflect our knowledge of the nature of language abilities and language use. Both the development and selection of authentic language tests is thus based on a theoretical framework that includes the language abilities of the test taker and the characteristics of the testing context.

Whether or not the framework one adopts is that proposed by Bachman (1990), or some other model, the implications for developers of tests and assessment tasks are clear: some kind of conceptualization of communicative language ability is needed which can serve as a starting point for deciding which abilities to sample and which test methods are most appropriate to tap these abilities. Thus in an oral test we would need to start with an idea of what we understand by 'speaking ability', what its components are, and which of these components we want to tap, drawing on the best of our knowledge of the nature of the ability or abilities in question. This would lead to a kind of sampling frame which enabled us to see which ability components were being sampled by which tasks using which methods. Shohamy (1993) provides an example of how this might be done in the development of a testing program used to provide diagnostic feedback to learners in Hebrew programs in the United States and Canada. Each of the sub-tests is based on current theories and understandings of the skill concerned and encompasses a wide range of tasks, text-types and item formats. Another example of this broad sampling approach is *Access:* a recently developed Australian Government English language proficiency test for prospective immigrants in which the Oral Interaction sub-test incorporates a variety of task types, topics and language functions (Wigglesworth and O'Loughlin 1993).

Summing up this discussion on authenticity, then, I would argue, along with Bachman and others, that authenticity in the sense of surface resemblance to target language use situations is a necessary but not sufficient condition for test and task validity. In addition, we need a principled way of specifying the abilities that assessment tasks are tapping and a sampling frame which enables us to obtain a complex and multidimensional picture of the way in which these abilities are being assessed through a range of different tasks using different assessment methods. One of the advantages of this multidimensional approach in the context of the classroom is that it naturally lends itself to profiling approaches which enable teachers to build up samples of different types of student work which reflect progress over a period of time.

3.2 Defining assessment criteria

The use of a wide variety of task-types places the onus on test developers (or teachers if they are responsible for assessment) to specify the characteristics of tasks with sufficient precision that they can be assessed. This means not only describing the tasks but also identifying a set of key criteria according to which learners' performance can be rated or scored. A range of different approaches has been adopted to task specification and identification of rating criteria. In this section I want to briefly discuss some of these approaches and to suggest ways in which data from studies of second language acquisition and use might be drawn on to inform task descriptors and rating criteria.

3.2.1 'Expert judgement' approaches

One commonly used way of producing criteria for proficiency testing is to ask expert judges to identify and sometimes to weight the key features of learner performance which are to be assessed. Experienced teachers tend to be the audience most frequently consulted in the development and refinement of criteria and performance descriptions (eg. Griffin and Nix 1991; Griffin and McKay 1992). In some cases they may be asked to generate the descriptors themselves by describing key indicators of performance at different levels of proficiency. In others, test developers may solicit comments and suggestions from teachers for modification of existing descriptors on the basis of their knowledge and experience.

The idea of using teachers' expert judgement appeals to logic and common sense. However it also brings with it certain difficulties. The first of these is that teachers' observations of language are bound to be influenced by the personalized constructs of language ability with which they operate. This is recognized by Griffin and McKay (1992:20), who adopted what they refer to as a 'bottom up consultative approach' to develop scales of ESL development for primary and secondary education in Australia and who write that:

> Limitations of this approach include the difficulties involved in obtaining appropriate descriptions of language behaviour from practitioners. It is often the case that practitioners' observations are limited by a lack of knowledge of theoretical models, by inadequate observation skills and/or an inability to articulate descriptions of independent student language behaviour. The developer of the scales has to make decisions about the need to use the imprecise language of the practitioner, and perhaps lose some of the definitive nature of the theoretical model, or to use a specialist terminology and run the risk of practitioner misinterpretation and rejection.

If expert opinion is to have any currency as a method of developing criteria, then one would expect that a given group of expert judges would concur, first on the criteria which make up the behavioral domain being assessed and second, on the allocation of particular performance features to particular levels. (Obtaining data in this way would be an integral part of construct validation). One would also expect that the group would be able to agree on the extent to which a test item was testing a particular skill and the level of difficulty represented by the item. (Agreement would constitute evidence for content validity).

Studies aimed at investigating how expert judgements are made, however, cast some doubt on the ability of expert judges to agree on any of these issues. Alderson and Lukmani (1989), for example, in an examination of item content in EFL reading tests, found that judges were unable to agree not only on what particular items were testing but also on the level of difficulty of items or skills and the assignment of these to a particular level. Devenney (1989) who investigated the evaluative judgements of ESL teachers and students of ESL compositions, found both within-group and between-group differences in the criteria which were used. He comments:

> Implicit in the notion of interpretive communities are these assumptions: (1) a clear set of shared evaluative criteria exists, and (2) it will be used by members of the interpretive community to respond to text. Yet this did not prove to be the case for either ESL teachers or students

On the basis of findings such as these it would clearly not be advisable to rely solely on teachers' expert judgement as a basis for determining assessment criteria. This is not to suggest that teachers do not have a role in identifying behavioral indicators and tasks that will be used to assess their students. However, as Griffin and McKay (1992:21) note, the data they provide needs to be cross-checked against theoretical research and other published data.

3.2.2 Rating scales

Another - and possibly the easiest - way to define criteria and descriptors for language assessment is to use those already in existence. There is no shortage of models and examples. Literally thousands of rating scales, band scales and performance descriptors are used throughout the world to describe aspects of language performance in a global way. These are frequently used as generalized criteria against which task performance can be rated.

A number of objections, have been raised, however, to some of the more commonly used proficiency rating scales, such as the ACTFL scale, used to certify foreign language teachers in the United States. These are discussed in detail by Brindley (1991) and North (1993) and will not be reiterated in detail here. However, in the context of TCA some of the more pertinent objections could be summarized as follows:

- The scales are not based on studies of second language use and as such, have no empirical support (Lantolf and Frawley 1988).
- The logic of the way levels are arrived at is essentially circular-'the criteria are the levels and vice-versa' (Lantolf and Frawley 1985:340). They cannot therefore be criterion-referenced in the accepted sense since there is no external standard against which the testee's behaviour may be compared.
- It is very difficult to specify relative degrees of mastery of a particular skill with sufficient precision to distinguish clearly between levels. This is

illustrated by Alderson's (1991:81-2) comment on the development of the IELTS Speaking scales:

> For some criteria, for example pronunciation or grammatical accuracy, the difference in levels came down to a different choice of quantifiers and we were faced with issues like is 'some' more than 'a few' but fewer than 'several' or considerable' or 'many'. How many is 'many'?

In addition to these problems, one of the major shortcomings in using generalized rating criteria is that they are too general to be applied to particular tasks. They do not, in other words, describe the qualities of individual task performances, nor do they describe what constitutes an acceptable standard of performance which is what is required in TCA (Pollitt 1991). Thus though rating scales of the general kind may be helpful in providing broad information for reporting purposes, they are of less use in assisting raters to make judgements relative to particular tasks and their construct validity continues to be surrounded by doubt.

3.2.3 Genre-Based Approaches to TCA

One way of obtaining detailed assessment information at the level of the individual task is represented by genre-based approaches to assessment which derive from the analysis of spoken and written genres within the framework of systemic-functional linguistic theory (Halliday 1985). Within this approach, the genres (such as *argument, describing a procedure*, etc.) are carefully described in terms of their structural organization and linguistic features. These features are then used as the basis for the implementation of a teaching-learning cycle and also serve as the criteria for assessment of overall task performance. In the context of primary English as a Second Language teaching, teachers have reportedly found this way of specifying tasks useful since it links the assessment criteria directly to what is being taught and focuses their attention on ways in which students are learning to make meaning (Mincham 1992). At the same time it offers the opportunity for learners to obtain diagnostic feedback on the extent to which they have met the criteria in the task, since the assessment checklist allows for differing levels of achievement to be recorded, ranging from 'very competent' to 'not yet'.

Genre-based approaches offer a way of describing and assessing language task performance that is underpinned by a powerful linguistic theory. Descriptions of genres provide explicit and testable hypotheses concerning the language demands of different text-types. The first step in a potentially valuable research agenda would be to determine to what extent the absence or presence of particular linguistic features in a text can be related to task difficulty. If a systematic relationship were shown to exist (as evidence from studies by Shohamy and Inbar (1991) and Pollitt and Hutchinson (1987) would seem to suggest), such information could be of great assistance in informing the rather vague descriptions of task and text characteristics which are used to rate task performance ('can understand more abstract texts', 'can make simple requests', etc.).

However, as Murray (forthcoming) points out, the genre approach relies on the availability of very comprehensive descriptions of different oral and written genres and 'full descriptions of the structures of most oral and written genres have yet to be developed.' From the practical point of view the amount of work involved in filling in individual checklists for large numbers of students could also prove quite daunting for teachers.

A variant on this approach is found in the Certificate in Spoken and Written English (Hagan et al. 1993), a competency-based curriculum framework used within the Australian Adult Migrant English Program for immigrant learners of English. The Certificate sets out outcome statements in the form of language competency specifications which describe the elements of the language performance in question, the criteria by which the performance is to be judged and the range of variables which obtain in the assessment situation (e.g. the amount of assistance the learner may receive). Though underpinned by the same systemic-functional theory of language, what distinguishes the assessment system used in the Certificate from the approach previously described is that here the performance criteria are mandatory. Before they can be awarded the competency, learners must demonstrate evidence of each of the performance criteria. Thus for the competency 'can negotiate complex/problematic exchanges', the performance criteria to be demonstrated are as follows:

> **Competency 4. Can negotiate complex/problematic spoken exchanges for personal business and community purposes**
>
> Achieves purpose of exchange and provides all essential information accurately.
>
> Uses appropriate staging for text, e.g. opening and closing strategies.
>
> Provides and requests information as required.
>
> Provides and requests goods and services as required.
>
> Explains circumstances, causes, consequences and proposes solutions as required.
>
> Sustains dialogue, e.g. using feedback, turn taking, seeking clarification and understands statements and requests of the interlocutor. (Hagan et al. 1993:76)

It is interesting to contrast this set of performance criteria with the following example taken from the Royal Society of Arts Practical Skills Profile Scheme which provides a method for assessing work-related and non-vocational courses in Communication, Numeracy and Process Skills.

> **Profile Sentence : C12 Participate effectively in negotiation**
>
> *Performance criteria (the student has demonstrated the ability to:)*
>
> Define own preferred outcome in given situation.
>
> State own needs/wishes clearly in language appropriate to listeners.
>
> Express disagreement sympathetically.
>
> Consider sympathetically suggestions of others.
>
> Suggest new ideas to solve temporary difficulties.
>
> Contribute to discussion freely and clearly without dominating the meeting.
> (Royal Society of Arts 1987:41)

This comparison illustrates graphically how the perspective of the test designer can influence the way that tasks are described and hence assessed. In the first case, it is primarily language that is the object of assessment (staging of discourse, conversational strategies, information giving, etc.). In the second, it is communication skills, in a more general sense, and non-linguistic, social and affective factors (intentionality, sympathy, empathy) which have a much greater role.

This raises a fundamental question that needs to be asked about TCA:

> If task fulfilment is the principal criterion (as it usually is) for assessment, then to what extent should non-linguistic factors be taken into account? (For example, Clark and Scarino (1993:32) include as one of the generic criteria for judgement of performance in the Hong Kong Targets and Target Related Assessment (TTRA) curriculum 'effectiveness of the product in relation to the purpose and context expressed in the task').

It is interesting, and perhaps significant, to note in the context of this discussion that disciplines outside applied linguistics interpret 'communication' or 'communicative competence' quite differently and hence employ different criteria for assessment. Communication theorists, for example, accentuate criteria such as *empathy*, *behavioral flexibility* and *interaction management* (Wiemann and Backlund 1980) and emphasise the role of non-verbal aspects of communication. In other fields, such as organisational management, communicative ability is seen very much in terms of 'getting the job done' and the success of communication is thus judged primarily in relation to how well the outcomes are achieved rather than on specific linguistic features (Brindley 1989:122-23). McNamara (1990:32) makes this point in relation to doctor-patient communication, noting that in the medical profession 'there is a concern for the communication process in terms of its outcomes.' He comments (1990:47) that 'sociolinguistic approaches to 'communicative ability' are indeed narrow, and narrowly concerned with language rather than communicative behaviour as a whole.' If TCA is concerned with task outcomes, then perhaps it is time for language testers who are concerned with certifying people's ability to get things done with language to reconsider the position they have conventionally taken with respect to these factors-that they are not part of communicative competence and are therefore not the object of assessment.

3.2.4 Towards data-based assessment criteria

I want to conclude this section on criteria for assessment by making some suggestions as to how criteria might be developed that are more consistent with current understandings of language acquisition and use.

First we need to compare data derived from studies of language in use conducted within a variety of theoretical paradigms with the descriptions of language skills and abilities that are used for assessment. Such research is important since what little work has been done tends to indicate major discrepancies between what actually happens and what test developers think happens. For example, Fulcher (1987) demonstrates that the criteria for fluency used in the ELTS Interview Assessment Scale do not reflect what happens in real conversational exchanges and that native speakers would, in fact, not meet the criteria. He recommends that data drawn from discourse analysis should be used to inform the constructs used in tests of oral performance. In a similar vein, Chalhoub-Deville (1993:20) in a study of the rating patterns of three groups of native speakers of Arabic assessing learners' oral performance concludes that 'research on L2 oral performance is needed ... that derives scales empirically according to the given tasks and audiences.'

At the same time, a lot more information is also needed about the cognitive demands that different types of tasks make on learners. In this regard, research evidence suggests that tasks that may appear to be of similar overall complexity may make different processing demands on learners (Bialystok 1991:121).

In the light of this and similar findings (see, for example, Snow et al. 1991; Chalhoub-Deville 1993), it is important to establish a principled way of describing and evaluating tasks. In order to throw more light on the question of task demands, as Bialystok (1991) points out, it will be necessary to undertake task analysis in different situations of language use. Bialystok suggests that task demands can be described in terms of two processing dimensions, namely *analysis of linguistic knowledge* and *control of processing*:

> Thus, the demands imposed upon language learners by various language uses can be described more specifically in terms of the demands placed upon each of these skill components, and the proficiency of learners can be described more specifically by reference to their mastery of each of the skill components (Bialystok 1991:64).

Once the qualitative characteristics of the tasks are known, banks of tasks can be developed and trialed. Using Rasch techniques, the tasks can then be calibrated on a common scale and related to defined bench-mark levels of performance. Griffin and McKay (1992) provide a description of how this can be done in a principled way.

Second, if assessment tasks are going to be closely related to 'real-world' tasks, then it is necessary to gather more information on those criteria that are used by other people outside the language classroom to make judgements on learners' task performance (Brindley 1991). This is particularly important in the case of learners in second language contexts. After all it is not teachers' judgements of students' language ability that will decide whether they manage to communicate in the 'real world.' Outside the classroom it is the lay person's impression of people's communicative effectiveness that will determine the extent to which learners' communicative goals are achieved. There is increasing evidence to suggest that non-teacher native speakers use quite different criteria in judging language performance from those used by teachers (Shohamy et al. 1992: Chalhoub-Deville 1993). If it is the judgement of these non-teachers that determines learners' communicative acceptability, it is necessary to investigate how these judgements are made, what criteria are used, and perhaps attempt to take these criteria into account in the construction of the instruments that are used to assess proficiency and achievement.

4. Reliability in TCA

4.1 The problem of human judgement

TCA relies heavily on teachers' subjective judgements of language performance. In the interests of fairness to learners, it is important that these judgements are seen to be reliable. As more and more rating tools are developed to assess productive task performance, teachers will need to be trained to interpret and apply assessment instruments in a consistent way. Rater training involving familiarization with the rating criteria and practice in applying them to samples of performances across a range of ability levels has long been standard practice with proficiency rating scales and it has been claimed that high levels of inter-rater agreement can be obtained in this way (e.g. Dandonoli and Henning 1991).

However the feasibility of obtaining inter-rater reliability with respect to language performance has come increasingly under question. Research in language testing has shown that despite training, 'significant and substantial differences between raters persist' and that rater behaviour can change significantly over time (Lumley and McNamara 1993). North (1993:45) in a comprehensive survey of the whole field of subjective judgements in rating concludes that 'judge severity is relatively impervious to training and that people rate in different ways.' This, of course, is hardly surprising given the complexity of the interaction between the language behaviour being rated, the personal characteristics of both the rater and the candidate, and aspects of the setting in which the rating takes place. However, it leaves the language tester in a dilemma: if variability in rater behaviour is the norm, then what - if anything - can be done to reduce the error in rater judgements?

4.2 The promise of new measurement technology

One possible solution to this problem is offered by recent advances in measurement technology in the form of multi-faceted Rasch analysis and its accompanying software package known as FACETS (Linacre 1988).

The Rasch model is one of a family of techniques known as latent trait theory or item response theory (IRT) which have been developed by psychometricians over the last three decades or so. One of the strengths of the theory is that it allows candidate ability and item difficulty to be estimated independently and reported on a common scale, thus avoiding many of the problems associated with sample-dependent classical measurement techniques (Henning 1987). The multi-faceted Rasch model extends previous Rasch models to include rater characteristics. It provides an estimate of candidates' ability based on the probability of a candidate obtaining a particular score on a particular task given the ability of the candidate, the difficulty of the item (in the case of language assessment this might be the rating category such as *fluency* or *cohesion*), the harshness of the rater, and the effect of any additional facets (Linacre 1989). The program adjusts candidate ability estimates to take account of raters' tendency to rate either harshly or leniently.

The use of FACETS can assist in the analysis of ratings of task performance in a number of ways:

- Since FACETS accepts variability and compensates for rater severity, it is not necessary to try to achieve complete agreement between raters. As long as raters are internally consistent, there is no need for raters whose rating patterns appear to be deviant to be excluded.

- It enables reports to be provided to raters showing their tendency towards severity and leniency. It also shows how each rater is using the steps on the scale.

- Through a technique known as bias analysis, it enables the interactions between different aspects or 'facets' of the rating situation to be modelled and examined, e.g. it is possible to see whether a certain rater is rating more or less harshly on a particular task or rating category (Lumley and McNamara 1993).

- It enables the rating categories used in assessing oral or written task performance to be subjected to scrutiny. If a rating category does not fit the underlying model, indicating an inconsistent pattern in scoring, it is flagged by the program as 'misfitting'. This allows rating criteria to be monitored and revised as necessary (Wigglesworth and O'Loughlin 1993).

I am not suggesting that new measurement technology can answer all the problems that will inevitably arise on the ground from the fact that raters will not always agree on the quality of task performance. It would be unrealistic to expect that many educational institutions would invest in the necessary training and expense associated with multi-faceted Rasch analysis. However there would appear to be no reason why co-operative research ventures could not be undertaken between educational institutions wishing to monitor the way in which subjective judgements are being made on task performance and on institutions with the necessary expertise in the use of the technology.

On a day-to-day level, institutions will still have to find ways of trying to achieve a common understanding and definition of different standards of learner performance. In this regard the collection and analysis of 'bench-mark' performance samples, accompanied by regular moderation sessions, has proved a useful way of focusing raters' attention on key aspects of task performance at different levels (Griffin and McKay 1992). Another way of trying to accommodate for rater severity without the benefit of technology is outlined by North (1993:45). He describes a procedure for oral assessment using two assessors, one who knows the class in question (high sensitivity) and one who is familiar with the whole range of the level (low sensitivity). Ratings are carried out independently using both holistic and analytical marking 'with negotiation over grades between the two assessors as a final step to adjust for severity' (ibid).

5. Practicality

Finally I would like to turn to the rather crucial issue of practicality. Though it is widely agreed that TCA has significant benefits, it is also likely to be affected by a number of pressures and constraints, including financial cost, time, expertise and demands for external accountability. To what extent can TCA be made to work, given these constraints?

5.1 Financial cost/time

There is no doubt that TCA is extremely time-consuming and by extension, expensive. Eliciting individual performances is much more difficult and time-intensive than administering pencil-and-paper tests. Commenting on the introduction of performance assessment in general education in the United States, O'Neil (1992:18) reports that 'some experts say performance assessments are likely to be at least two or three times more expensive per student.' Worthen (1992:452) suggests that 'the labour intensity of scoring and the need to observe performance over extended periods are primarily responsible for the high costs of performance assessment.'

Nuttall (1992:56) notes that teachers who administered the 1991 Standard Assessment Tasks in the UK, although they had learned new things about children's attainment 'found the tasks to be demanding of their time and energies; invariably, to prepare, administer and grade them required an average of 44 hours.'

On a similar note, Barrs (1992:55) comments that a common concern voiced about the implementation of the detailed observational recording system used with the Primary Language Record in the UK was the sheer amount of time necessary to document many student performances on an ongoing basis:

> Keeping detailed observational records of up to thirty children seems just too difficult.

She observes, however, that this aspect gradually became more manageable but noted that:

> ... it does seem to be the case that it takes a full school year to "learn the forms", to internalise the ways of observing that they encapsulate and to see the full value of this kind of recording (Barrs 1992:56).

These experiences would indicate that TCA has to be seen as a long-term rather than a short-term investment.

5.2 Teacher development

TCA is demanding not only in terms of time but also of teacher skill and a considerable investment in teacher development is necessary if teachers and learners are to obtain the maximum benefit from its use. In this context it should be noted that changing teaching and assessment practices or adopting new tools is no different to introducing a new curriculum or a new textbook. The introduction of TCA is an exercise in change management which by definition means trying to plan for the implementation of whatever change is proposed as a result of the professional development activities that are offered. If educational administrators are concerned with the long-term effects of what they do, then they need to be aware of this. Workshop participants who return to their institutions full of enthusiasm for new assessment ideas or tools cannot automatically be expected to apply their ideas. If they are to change their assessment practices or systems, they require support in terms of time, funding, resources and sometimes skilled support personnel (Fullan 1982).

The importance of providing this support cannot be overestimated. If teachers are not given adequate assistance in understanding and implementing new modes of assessment, the whole purpose of the introduction of TCA may be undermined. This is graphically illustrated by Shohamy (1993) in a study of the impact of three different types of language tests on teaching and learning in Israel. She reports on the introduction of a form of TCA in the Israel-EFL oral test, which is part of the national matriculation examination taken by high school students at the end of twelfth grade. Although the rationale of the test was to place greater emphasis on oral proficiency and improve students' speaking skills, she found that teachers perceived oral language 'exclusively in terms of testlike activities.' Thus when asked to define 'oral language' teachers often gave answers such as 'It is a role play' or 'It is an interview'.' She concludes (1993:15) that 'in terms of the nature of the test effect, in all three cases the results showed the instruction became more testlike and that this was most likely a result of teachers not having been trained to teach the new areas being tested' and adds that 'when teaching and testing become synonymous, the tests become the new de facto curriculum.'

5.3 External accountability

One of the main difficulties in introducing TCA in general education has been its public acceptability. Nuttall (1992) reports that the introduction of performance assessment into the school curriculum in the UK has alarmed some people who are used to standard pencil-and-paper assessment and who feel that the new kinds of assessment are less rigorous:

Unfortunately the minority who take this view have the ear of Prime Minister John Major, who has decided that the amount of performance assessment must be significantly curtailed. His exact words were:

> By testing I do not mean some weird experiment in a corner. (i.e. a reference to the Standard Assessment Tasks). What I mean is pencil-and-paper testing for a classroom so people have a measure of how they are doing-see if there is a problem so that you can put it right.

Nuttall (1992:57) reports that the day after the Prime Minister's pronouncement, the development contracts for Grade 2 Standard Assessment Tasks were cancelled and new tenders called for the development of pencil-and-paper tests.

Hopefully the justification for the use of TCA in language assessment is more self-evident. However Nuttall's comments are a sobering reminder that whenever new forms of assessment are introduced they need to be accompanied by very clearly presented and accessible statements explaining their purpose, use and justifying their financial costs. The latter is a particularly important factor where large-scale oral assessment is concerned.

6. Conclusion

In this paper I have identified a number of key issues relating to the validity, reliability and practicality of task-centred assessment and I have suggested ways in which some of the potential difficulties associated with its implementation and use might be addressed.

Overall, experience in general education in a number of countries seems to indicate that TCA can be made to work. This is, I suspect, because it has the strong support of both teachers and learners - it takes assessment out of the realm of something which is done *to* students into the realm of something that can be done *with* them. In the words of Broadfoot:

> In place of ubiquitous competition and external judgement, assessment is harnessed to teaching to provide explicit statements of curriculum goals; to equip pupils with the skills to set their own goals and review progress towards them; to make pupils jointly responsible with teachers for both formative and summative reviewing and reporting (Broadfoot 1988:5).

However, a number of challenges remain. In the first place, in relation to the validity question, as I have indicated at various points throughout this paper, a lot more work needs to be done in order to develop assessment criteria which reflect current theories of language learning and language use. Second, as far as reliability is concerned, if TCA is to have public credibility, the problem remains of trying to ensure consistency in the application of assessment criteria. While new measurement technology may provide some solutions to this problem, this technology will only be available to a few. As some testers have suggested, one of the consequences of adopting TCA may well be learning to rethink the notion of reliability:

> ... if we indeed value clinical judgement and a diversity of opinions among appraisers (such as certainly occurs in professional settings or post-secondary education), we will have to revise our notions of high-agreement reliability as a cardinal symptom of a useful and viable approach to scoring student performance. We will have to find a previously uncharted course between insisting on uniform judgements and mayhem. Possibly, we will have to seek other sorts of evidence that responsible judgement is unfolding-that participants agree on the relevant categories for describing performance, that scores fall within a certain range, or that recipients can make thoughtful use of the range of opinions offered to them (Wolf et al. 1991:63).

Finally, at a practical level, the adoption of TCA has a number of major consequences at all levels of an educational system. For teachers and learners it means that they will need to become accustomed to thinking of language tasks not only as activities but also as indicators of progress and achievement. Learners will thus need to understand the criteria according to which their performances will be judged. This, in turn, will necessitate a closer examination of the components of language tasks and a raising of learners' awareness of how language functions to achieve particular communicative purposes. As far as 'consumers' of assessment information are concerned, the reporting of task performance means that they may have to be persuaded to accept assessments that are complex, qualitative and multidimensional, rather than uniform and standardized. This will not be easy and will necessitate close co-operation and continuing dialogue between all of the stakeholders involved in language programs.

There is no doubt that task-centred assessment in language learning is firmly established at the level of the classroom where it has demonstrated the potential to bring about significant improvements in the quality of learning (Shohamy 1993). As it moves into high stakes areas such as certification and selection, however, it remains to be seen whether the momentum will continue. In this regard, we can only hope that the value of TCA will become as evident to those outside the classroom as it is to those within it.

References

Alderson, J.C. & Lukmani, Y. (1989). Cognition and reading: cognitive levels as embodied in test questions. *Reading in a Foreign Language*, 5(2).

Alderson, J.C. (1991). Bands and scores. In J.C. Alderson & B. North (eds.), *Language Testing in the 1990s*. London: Macmillan.

Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Barrs, M. (1992). The primary language record: what we are learning in the U.K. In C. Bouffler, (ed.) *Literacy evaluation*. Sydney: Primary English Teaching Association.

Bialystok, E. (1991). Achieving proficiency in a second language: A processing description. In R. Phillipson, R. Kellerman, L. Selinker & M. Sharwood Smith, M. (eds.) *Foreign / Second language pedagogy research*. Clevedon, Avon: Multilingual Matters.

Broadfoot, P. (1988). *Introducing profiling*. London: Macmillan.

Breen, M. P. (1987). Learner contributions to task design. In C.N. Candlin & D. Murphy (eds.) *Language learning tasks*. Englewood Cliffs: Prentice Hall.

Brindley, G. (1986). *The assessment of second language proficiency: Issues and approaches*. Adelaide: National Curriculum Resource Centre.

Brindley, G. (1989). *Assessing achievement in the learner-centred curriculum*. Sydney: National Centre for English Language Teaching and Research.

Brindley, G. (1991). Defining language ability: the criteria for criteria. In S. Anivan (ed.) *Current developments in language testing*. Singapore: Regional Language Centre.

Broadfoot, P. (1988). *Introducing profiling*. London: Macmillan.

Chalhoub-Deville, M. (1993). Performance assessment and the components of the oral construct across different tasks and rater groups. Paper presented at Language Testing Research Colloquium, Arnhem, August.

Clark, J. & Scarino, A. (1993). The integration of language and content in TTRA: implications for schools. In N. Bird, J. Harris & M. Ingham (eds.) *Language and content*. Hong Kong: Government Printer.

Crookes, G. (1986). *Task classification: A cross-disciplinary review*. University of Hawaii at Manoa: Centre for Second Language Classroom Research, Social Science Research Institute.

Dandonoli, P. & Henning, G. (1991). An investigation of the construct validity of the ACTFL proficiency guidelines and oral interview procedure. *Foreign Language Annals*, 23(1):11-22.

Devenney, R. (1989). How ESL teachers and peers evaluate and respond to student writing. *RELC Journal*, 20(1).

Fulcher, G. (1987). Tests of oral performance: The need for data-based criteria. *ELTJ* 41(4).

Fullan, M. (1982). *The meaning of educational change*. Toronto: Ontario Institute for Studies in Education.

Griffin, P. & Nix, P. (1991). *Educational assessment and reporting*. Sydney: Harcourt Brace Jovanovich.

Griffin, P., & McKay, P. (1992). Assessment and reporting in ESL Language and Literacy in Schools project. In P. McKay (ed.). *ESL development: Language and literacy in schools. Volume II: Documents on bandscale development and* *language acquisition*. Melbourne: National Languages and Literacy Institute of Australia.

Hagan, P., Hood, S., Jackson, E., Jones, M., Joyce, H. & Manidis, M. (1993). *Certificate in spoken and written English.* Sydney: NSW Adult Migrant English Service and the National Centre for English Language Teaching and Research.

Halliday, M.A.K. (1985). *An introduction to functional grammar*. London: Edward Arnold.

Henning, G. (1987). *A guide to language testing*. Cambridge, MA: Newbury House.

Lantolf, J. & W. Frawley (1985). Oral proficiency testing: a critical analysis. *Modern Language Journal*, 69(2).

Lantolf, J. & W. Frawley (1988). Proficiency: defining the construct. *Studies in Second Language Acquisition*, 10(2).

Linacre, J. M. (1988). *FACETS: A computer program for the analysis of multi-faceted data*. Chicago: MESA Press.

Linacre, J. M. (1989). *Multi-faceted measurement*. Chicago: MESA Press.

Long, M. H. & G. Crookes (eds.) (1993). *Tasks in a pedagogical context*. Clevedon, Avon: Multilingual Matters.

Long, M.H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (eds.) *Modelling and assessing second language acquisition*. Clevedon, Avon: Multilingual Matters.

Lumley, T. & McNamara T.F. (1993). Rater characteristics and rater bias: implications for training. Paper presented at Language Testing Research Colloquium, University of Cambridge, August 1-2.

McNamara, T.F. (1990). Assessing the language proficiency of overseas-qualified health professionals. Ph. D. Thesis, University of Melbourne.

Mendelsohn, D. (1989). Testing should reflect teaching. *TESL Canada Journal*, 7(1).

Mincham, L. (1992). Assessing the English language needs of ESL students. In C. Bouffler (ed.) *Literacy evaluation*. Sydney: Primary English Teaching Asociation.

Murray, D. (forthcoming). Using portfolios to assess writing. To appear in *Prospect*, 9(2).

North, B. (1993). *The development of descriptors on scales of language proficiency*. Washington, D.C: The National Foreign Language Center.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.

Nunan, D. (1993). Task-based syllabus design: selecting, grading and sequencing tasks. In M.H. Long & G. Crookes (eds.).

Nuttall, D. (1992). Performance assessment: the message from England. *Educational Leadership*, May.

O'Neil, J. (1992). Putting performance assessment to the test. *Educational Leadership*, May.

Pienemann, M., Johnston, M. & Brindley, G. (1988). Constructing an acquisition-based assessment procedure. *Studies in Second Language Acquisition*, 10(2).

Pollitt, A. (1991). Response to Charles Alderson's paper: 'Bands and scores'. In J.C. Alderson & B. North (eds.) *Language testing in the 1990s*. London: Macmillan, pp. 87-94.

Pollitt, A. & Hutchinson, C. (1987). Calibrating graded assessments: Rasch partial credit analysis of performance in writing. *Language Testing* 4(1):72-92.

Royal Society of Arts (1987). *Practical skills profile scheme*. London: Royal Society of Arts.

Shohamy, E. (1992). New modes of assessment: the connection between testing and learning. In E. Shohamy & R. Walton (eds.). *Language assessment for feedback: Testing and other strategies*. Dubuque, Iowa: Kendall/Hunt Publishing Company.

Shohamy, E. (1993). *The power of tests: The impact of language tests on teaching and learning*. Washington, D.C.: The National Foreign Language Center.

Shohamy, E. & Inbar, O. (1991). Validation of listening comprehension tests: the effect of text and question type. *Language Testing*, 8(1):23-40.

Shohamy, E., Gordon, C. & Kraemer R. (1992). The effect of raters' background and training on the reliability of direct writing tests. *Modern Language Journal*, 76(2).

Skehan, P. (1984). Issues in the testing of English for specific purposes. *Language Testing*, 1(2).

Snow, C.E., Cancino, H., de Temple, J. & Schley, S. (1991). Giving formal definitions: a linguistic or metalinguistic skill. In E. Bialystok (ed.) *Language* *processing in bilingual children*. Cambridge: Cambridge University Press. pp. 90-112.

Spolsky, B. (1985). The limits of authenticity in language testing. *Language Testing*, 2(1).

Swain, M. (1993). Second language testing and second language acquisition: Is there a conflict with traditional psychometrics? *Language Testing* 10(2).

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Wigglesworth, G., & O'Loughlin, K. (1993). An investigation into the comparability of direct and semi-direct versions of an oral interaction test in English. *Melbourne Papers in Language Testing*, 2(1).

Wiemann, J, & Backlund, P. (1980). Communicative competence. *Review of Educational Research*, 50(1).

Wolf, D., Bixby, J., Glenn, J. & Gardner, H. (1991). To use their minds well: New forms of student assessment. *Review of Research in Education* 17.

Worthen B. (1993). Critical issues that will determine the future of alternative assessment. *Phi Delta Kappa*, February.

PART TWO

TASK-CENTRED LEARNING AND ASSESSEMENT

INTRODUCTION

Task-based teaching and learning is part of a current trend in language teaching circles. In Hong Kong, the recent attempt to introduce a curriculum based on Target and Target Related Assessment (now modified to Targets Oriented Curriculum) has resulted in heated debate in education circles as the theoretical underpinning and the implementation of this curriculum innovation is debated.

The first paper in this sub-theme is a broad survey article by *Stephen Hall.* His comprehensive coverage of the research literature provides a welcome introduction to the field.

The second paper by *David Nunan* could have appeared in any of the sub-themes of the Conference, but was included here because of the task-based nature of the writing projects he describes and because of the functionally-based approach he takes to the description of the texts described in the paper. In quoting Halliday (1995) who says that every text unfolds in some context of use, ***Nunan***, using genre as the basis for his approach, states that the typical rhetorical structures and grammatical forms which reflect the communicative purpose or intent of the genre in question allows for that genre to be analysed and evaluated. ***Nunan*** asserts that this mapping of function, structure and form onto text allows for a task-based approach to assessment.

Colin Barron describes in the third paper the conflict between the alteric nature of the Spack (1988) approach to ESP and the nature of the tasks which students should be set and the mimetic model described by Widdowson (1983). Instead of taking up a middle position, ***Barron***, in a challenging paper, tries to place ESP "firmly within the multidimensional space that constitutes the students' disciplinary culture", and exemplifies this approach in the context of a task-based assignment set for his students.

The fourth paper contains a description, by *Brenda Bushell* and *Brenda Dyer* of a task-based approach to curriculum development and course design practised at the International Christian University, Tokyo. This was one of many papers presented at the conference which describes current attempts to initiate a task-based approach in the classroom and was selected for publication because of the interesting way in which the authors attempt to set the tasks within the context of global education.

The final paper by *Gertrude Tinker Sachs, Stella Kong, Anne Lo* and *Tom Lee* explores teachers' basic understanding of a task description to ascertain what teachers would do if asked to teach from a description. It examines the extent to which primary and secondary teachers differ in their conceptualizations and interpretations. This is particularly apposite in Hong Kong at present because of the controversy described above, created by the government's decision to introduce a task-based curriculum into primary and secondary schools in September 1994.

Peter Falvey

TASK AS A UNIT OF TEACHING ANALYSIS

Stephen Hall

Introduction

The common student statement of having "work" to do, is a valid metaphor for the language learning process. Tasks have to be mastered and goals met for the learner to gain fluency and confidence. If the learning burden is too great the student will not benefit as fully as the teacher may anticipate (Nation 1983). The work or task will run the risk of being an activity where student perceptions of what has to be done do not correlate with teacher expectations (Peterson et al. 1982). The question of what is really happening as a student learns can be examined by applying the concept of "task" to classroom processes. This can be done by defining task as a unit of teaching analysis, considering the role of the product in a classroom, and examining the influence of processing conditions on task-based learning. Task learning arrangements are also a useful area for analyzing what happens in a classroom. A classification of tasks can also include a description of task types in terms of information processing.

Defining Task

The classroom question of "What do I do now?" is often heard in the outside world when work has to be done. The parallel construct of school work as a set of tasks that must be processed is used by researchers to describe classrooms. Doyle (1983) has been influential in analyzing school processes using work as the main metaphor, linking analysis of cognitive processes outside the classroom setting to the cognition of academic tasks.

With work as the framework he states:

"the term 'task' focuses attention on three aspects of students' work:

1. the products students are to formulate....

2. the operations that are to be used to generate the product

3. the "givens" or resources available to students while they are generating a product....Academic tasks, in other words, are defined by the answers students are required to produce and the routes that can be used to obtain these answers" (Doyle 1983:161).

Doyle's definition is an economical summation of the concepts of task. His model is paralleled in the second language field by Mohan's (1986) term "activity."

For Mohan "activity is a broad integrating idea relevant to all teaching and learning" (1986:45). This concept involves the linking of content and language, where he points out that communicative language teaching uses general background for the "specific practical side." There is a focus on integrating content and form.

Experience in developing a programme of wide ranging communicative language teaching procedures, known as "The Bangalore project", led to Prabhu's (1987) widely known definition of a task:

> An activity which required learners to arrive at an outcome from given information through some process of thought and which allowed teachers to control and regulate that process was regarded as a 'task' (1987:24).

Prabhu's classroom-experience based definition is similar to Doyle's model in that Prabhu (1987:24) defines task further into "two parallel tasks" and a third component:

1. Pre-task is in the form of a whole class preparatory activity.
2. Task itself is when the students do individual or group activity.
3. Marking and feedback is when concern is with the outcome of the task and the product.

The "givens" are provided, the operations undertaken and the product assessed.

A cross-disciplinary review of task by Crookes (1986:1) notes the developing utility of the term and for the purposes of his work he defines task as:

> A piece of work or an activity, usually with a specified objective, undertaken as part of an educational course, at work, or used to elicit data for research.

His first two definitional phrases, that is work with a specific objective, are identical to Long's definition of task as a unit of syllabus design (Long 1985:89). A particular objective, answer or outcome is central to Breen's (1987:23) definition of task as:

> any structural language learning endeavour which has a particular objective, appropriate content, a specified working procedure, and a range of outcomes for those who undertake the task. 'Task' is therefore assumed to refer to a range of work plans which have the overall purpose of facilitating language learning.

Richards (1990:39) in devoting a section of a chapter entitled "Beyond Methods" to "tasks," notes that tasks:

> refer to activities that teachers assign to attain particular learning objectives. For any given subject at any given level, a teacher uses a limited repertoire of tasks that essentially define the teacher's methodology of teaching.

This point that the tasks or the classroom work of a teacher, defines methodology will be developed further in this paper by reference to classroom-based research.

In recent work Long and Crookes (1992) broaden the claim that task is a unit of analysis that is applicable to all phases of course design. They also note that the emphasis is on something that is done, not something that is said. This emphasis can be seen in practice in a learner-led task based course called *Go for Gold*, a business English course run by the language and communication department at the Papua New Guinea University of Technology. It is a long-standing and successful course which is a working example of a task-based syllabus "designed on the premise of learning by doing" (Hyland and Hyland 1992).

The reality of these general definitions is a recognition that the work of a classroom, the obtaining of the "answer", is the central organising principle, rather than a teacher chosen decision of what language item is being taught. As early as 1979 Breen, Candlin and Waters (1979:56) pointed out that "it is at the level of tasks that the actual working process of the classroom group is realized in terms of what is overtly done from moment to moment within the classroom."

Task as a term can be applied to target tasks and pedagogic tasks and utilised to analyze the relevance of teaching for learner needs both within and outside the classroom. Nunan (1991:279) in describing what he and others term task based language teaching (TBLT) writes of the "attempt to link classroom language learning with language activation outside the classroom." This defining of the basic concept of task into target tasks and pedagogic tasks is elaborated later in this paper. It is a feature of many ESP programmes which recognise the need to align the language work of a classroom with the language work outside of the classroom. The aligning of the means that learners employ to get answers in the classroom with strategies needed in the community is an area which may well concern those in general English teaching, as well as those in the more specialised areas.

"Activity" or task, which Mohan, Prabhu and Crookes use interchangeably will be analyzed further and discussed in terms of the finished work, or the product. Secondly, the routes to the answer, the operations will be described. The influence of the task itself on learning and the conditions will be considered.

A focus on an analysis of teaching will be developed through discussion of effective task types. Task will be analyzed using Doyle's framework:

1. Product

2. Givens or Conditions.

3. Operations

Task and the Product.

As a result of task related efforts, students produce a finished piece of work. This is often the main focus for teacher, student and those to whom both are accountable. A student is often asked "Can I see the work you did in class today?" rather than "What did you learn?" The product is frequently the focus of a classroom, the point of application of a student's cognitive plans (Marx and Walsh 1988).

A concern for accountability continues to put emphasis on product with many teaching situations showing evidence of behavioral influences (Bloom 1976, Carter and Doyle 1982). Related techniques, such as mastery learning and programmed learning call for high rates of success and by their very design, focus on product (Berliner 1987). Constant success and large amounts of reinforcement can become the main motivating influence on the learning process with the primary focus being on the product. The process is one where "accountability drives the task system" (Doyle 1983:185), so that analysis of product for the form of work which it demands, becomes important. This is not to deny the much focussed on aspects of the process of learning but clarifies the importance of the "answers" and the everyday classroom reality of the product as an important determinant of work. This will be discussed further in terms of the task product and student perception and the role of the form of the product. Teacher viewpoints of what is going on in a classroom are often subordinate to student perception and how students see their work.

Product and Student Perception

When one focuses on student perception of a product one sees the validity of Doyle's model in that he writes of the "givens," the conditions that affect the product. In achieving the product, the conditions under which the task is done are a major influence on the quality of the product. Much has been written about the physical conditions of learning, yet student attitudes and perception are a critical condition in the quality of learning that is happening.

In analyzing the products of learning it is easy to overlook the students' perceptions of the product: that is, "What is this work for?" The concern is not always "What's it about?" but what value is attached to the task. The best teaching expectations, even in the name of learner-centered programmes can be overridden by student perceptions of how the product will be evaluated (Winne and Marx 1982). How many times as classroom teachers have we heard the question, "But how many marks is it worth? Will it count towards the final grade?"

There can also be differences between the lesson's cognitive objectives from a teacher's viewpoint and students' performance which is often related to student expectations of evaluation (Blumenfeld, Mergendoller and Swarthout 1987, Anderson, Stevens et al., 1988).

Product and Form

A second major aspect in considering the product as a dynamic in task process is the form of the product itself. The work to be done, that is the obtaining of an answer or product, will often override student concern with content (Mohan 1986; Prabhu 1987).

Blumenfeld, Mergendoller and Swarthout (1987:140), summarize this with "content drives cognition, the form drives behaviour." This is a point that some contemporary classroom material does little to include. That is, the form of the answer and the students expectations of how much work has to be done to get the answer will greatly influence the learning.

The form of the task itself is the mediator between teaching behaviours and learning behaviours. As such Bennett (1988), calls attention to understanding the form of tasks, rather than focusing on prescriptive programmes and overtly structured classrooms.

In cognitive psychology, Anderson (1980:144), among others, points to the importance of what is produced when performing a task.

> Continued exposure to tasks of low cognitive complexity and challenge in forms where the information is clearly framed and easily accessible, with products that require little transformation of information and allow for little definition by the student is likely to result in preferences for easy, clearly defined task forms which require minimal time or involvement on the part of the learner.

There are several reasons why the product of a lesson is an important part of task analysis. The product or answer to a teacher-set objective is a major concern for students, who are often concerned with the evaluation of a task. The answer and efforts to find it are related to the students' view of a task, to the extent that the task itself may be a mediator between students' performance and the teacher's expectation.

The form of the product may be considered as an important aspect through the influence of form on the acquisition of language. Few would advocate a return to *forms* per se as Long and Crookes point out (1992), but the influence of *form* in the processes of language acquisition is well documented. Long and Crookes (ibid.) cite work on marked and unmarked forms transferring to implied marked and unmarked forms (Zobl 1985) and describe the importance of attention to form in task planning, a somewhat self evident observation for those involved in everyday classroom work.

Classroom research in the context of immersion learning found that involving teaching of form and an analytic approach complements the experiential approach (Allen, Swain, Harley and Cummins 1990). The authors describe the experiential approach to teaching as being marked by "a combination of activities marked by group work, broad range of reference, use of extended text, (and) reaction to message rather than code" while classes with an analytic approach involve "relatively more time on whole class activities, form-focussed practice, use of minimal text, (and) reaction to code rather than the message" (ibid.:58). Swain and colleagues' study of eight French grade 11 immersion classes was based on classroom observation and addresses pedagogic issues of how attention to form affects outcomes of learning. A correlational analysis related a range of finely differentiated observational variables to learning outcomes or products. The results suggest a place for the analytic focus, that is a focus on form.

> The results of the correlational analysis suggest that core French students benefitted from a generally experiential approach in which relatively more time was devoted to such features as information gap, reaction to message, and topic incorporation. At the same time, there were positive correlations between various form-focused, teacher directed activities and adjusted post-test scores. These results lead us to the conclusion that the analytic focus and the experiential focus may be complementary, and that they may provide essential support for one another in the classroom. (ibid.:62)

A concern with form has led to discussion in the field of oracy research of the need for balance in tasks that develop accuracy, attention to code, as well as fluency (Murphy 1991). Fluency tasks need to be balanced by attention to accuracy and form as interaction-based tasks may not create negative input that shows a breakdown in communication and hence a realisation that a form is ungrammatical (White 1987).

Focus on form is useful in determining that learning is at the level of the learners. With appropriate levels of learning, students will pass through a developmental sequence and extend the scope of rules to more general fluency and production (Pienemann and Johnston 1987). Long and Crookes (1992) state that "the evidence does motivate a focus on form that is, use of pedagogic tasks and other methodological options which draw student's attention to aspects of the target code" (1992:43). Recent attention to integrating form and interaction can be seen in the recent theoretical writing and textbook production given to presenting grammar from a "new" perspective.

Concern with the answer or product in a learning session is relevant to the student's perception of a classroom, related to the evaluative process and also needs to be relevant, as form affects learning. The product or answer therefore needs to be well integrated with an assessment of students' abilities to attain the product and it needs to be linked to task design which considers conditions of learning.

Task and Conditions

Teacher centred and classroom environmental research suggests that the learning task and the conditions of the work are of importance in understanding classroom interaction (Richards and Rodgers 1986). Curriculum allocation, more time on task, or improvements in levels of student involvement would not seem fruitful, if the prime concern is not with the tasks themselves (Doyle 1979, Burns 1984).

Bennett (1988) has applied consideration of the conditions of a task to his influential work in the British educational system. He describes how learners cope with classroom activities:

> This perspective assumes that the tasks in which pupils engage, structure to a large extent what information is selected from the environment and how it is processed. Tasks organize experience and thus an understanding of that experience and the process of acquisition, first requires an understanding of the tasks on which pupils work. (ibid.:24)

Time on Task

The utility of task as a unit of analysis, can be considered in the light of research on time on task. The field, a readily quantifiable condition for learning, has been widely studied for a considerable period (Carroll 1963, Denham and Lieberman 1980, Berliner 1983, Gettinger 1984, Gettinger 1989).

It is found generally that time involved in an academic task affects performance, when measured in terms of academic success.

Studies do not, however, show that increasing time on task will in itself improve results (Frederick and Walberg 1980, Carroll 1985). Time is only one factor. In overviews of the work (Bennett 1982 1988) a wide variety of results are presented. Time on task is further analyzed into the time actually available and secondly into assessing involvement with content (Mohan 1979) and instruction. When one begins to assess involvement one begins to look at the product and the steps taken to get there; it is an analysis of the work itself, with the activity being of importance (Shavelson and Stern 1981, Burns 1984).

Although there are positive relationships between time and learning in Karweit's detailed reviews of studies (cited in Croll 1988), she emphasises that the task itself and the learning arrangements are critical factors. Gettinger's (1989) study of 118 third grade children reached similar conclusions by examining time spent on learning and time needed for learning. Results show that it is possible to increase student perseverance but that the "effects of an increase in time spent, however, are not consistent"(1989:89). She notes a need for research on time to focus on factors other than the effects of specific time periods on achievement. The lack of a clear relationship between time and achievement in her findings parallels other work which argues that it is the manner in which time is spent which is crucial to achievement (Anderson 1981, Wang 1987). Gettinger cites her work on individual learner differences (1984) to highlight the importance of factors other than time itself; that is the tasks themselves, modes of instruction, learning arrangements and importantly the learners. With the growing awareness of the need for learner-centred programmes it is worthwhile to return to learner perceptions of the role of time spent on tasks in the classroom. Any examination of work in a classroom needs to consider learner perceptions.

Time and Student Perception

Task analysis is focused on the learner if we accept that working with tasks and finding the routes to the answer are the basis of a classroom. We thus come back to the question of what the learner is doing. Peterson and colleagues (Peterson, Swing, Braverman and Wass 1982) raise questions about the thinking that students use while working. They researched task and student perception during mathematics instruction, and found that achievement correlated more highly with students' self reports of whether they were on task or not, than with classroom data of observable on-task behaviour.

Secondly, achievement correlated with the use of direct instruction and specific cognitive approaches such as semantic mapping (Stahl and Vancil 1986). The focus moves to the learner in that classroom wide activities were unrelated to achievement. These findings point to the importance of cognitive engagement created by the task as being of more importance than time on task per se. Secondly, Peterson and colleagues note that all classroom processes will involve the learners' perceptions as a vital condition for improving the quality of performance.

Classroom Conditions

Teaching Methodology

The classroom has often been seen as the laboratory for a particular method or teaching style, where the teacher creates the learning environment. Much work exists on what makes an effective teacher (Swaffar, Arens and Morgan 1982, Brophy and Good 1986, Doyle 1986).

Evidence of the usefulness of task as an analytical unit for effective teaching comes from the research of Brophy, Rohrkemper, Rashid and Goldberger (1983), who found that teachers' motivational techniques and concentration on content were unrelated to student engagement. Different methodologies did not correlate with work done.

In a major review of research into teacher planning and instructional modes, Shavelson and Stern (1981:447) found that a discrepancy exists between the way teachers are trained and what happens in the classroom.

> Research on teacher planning has found that the instructional activity is the basic unit of planning" (Clark and Yinger 1979, Peterson et al. 1978, Smith and Sendlebach 1979, Zahorik 1975).

Shavelson and Stern suggest that the environment of the classroom and the demands of management lead to decisions about academic work being of prime importance.

Swaffar, Arens and Morgan (1982) (cited in Crookes 1986), researched the relationship between methods or approaches and actual classroom practice. Teachers using supposedly different methods were found to use similar classroom practices. They found that labels for methodology are not useful. Gaies (1983) states that this finding is an important basis for process research. Bennett (1988:27) cites a recent House of Commons Select Committee Report and its call for a move from analysis of styles to a more valid model for learning and teaching. He concludes that:

> Research on opportunity to learn, in emphasizing the quantification of time, neglected to characterize the nature and quality of classroom tasks and in common with the teaching styles approach it neglected the process of learning itself.

He calls for a direction of research which Swaffer et al. state is important. The clear statement which concludes Swaffar et al.'s (1982:32) findings in independent support of Shavelson and Stern's study is that:

> any analysis of methodologies needs to commence in terms of task, order (of tasks) and learning strategies. This is the way we, as foreign language teachers, interpret the pragmatics of the classroom.

This interplay between tasks, task ordering, task types and learning strategies creates a classroom dynamic of conditions.

The Dynamic of Conditions

Certain dynamics of tasks in a classroom can be delineated as they provide insights into task design principles. A general summary of Doyle's (1983) use of task in assessing classroom dynamics follows:

1. Academic tasks organize students' information processing. Accountability drives the task system.
2. Answering is the task in classrooms. Attention goes to the answering event itself rather than simply to the content.
3. Higher level cognitive processes are difficult to accomplish in the classroom due to ambiguity, risk and class management problems.
4. The knowledge students have of content is "probably embedded in their cognitive representations of the tasks they encounter in classrooms." (Doyle 1983:186).
5. Due to group dynamics and often public evaluation teachers are under pressure to adjust tasks.
6. Group management and answering will often put the attention on getting the work done rather than on the quality of the work.
7. The amount of work and adaptation to tasks will create pressure to maintain stability in the task system. Familiarity and predictability simplify the task.
8. The classroom itself as a task environment creates difficulties for younger and less able students (ibid.:185-186).

This brief summary shows the utility of task. It is beyond the scope of this paper to focus on classroom management and physical conditions of learning. I shall, however, look at the role of learning arrangements, the influence of task types and student interaction with tasks.

Student and Task

Students are concerned with working through and completing classwork, or at the worst, devising means of avoiding completion if it is beyond their capabilities. Given that teacher expectation and students' viewpoints can mismatch, it becomes critical that tasks are matched with students' levels in terms of the work itself. The type of task will influence the benefits that can be gained in classroom learning. An analysis of tasks could aid in understanding how students can benefit from the classroom as teacher and students pursue answers.

Previous assumptions of the role of the student have seen them as moving through a lockstep progression in a curriculum or method which is externally prescribed. Bloom's (1956) taxonomy has been widely utilised in not only a descriptive manner.

The framework of setting objectives to move through a pre-determined hierarchy of skills can be compared with the frameworks of Cummins (1982) and Tikunoff (1985).

Cummins describes the common setting where a student needs to function at the teacher-designated level to achieve an instructional objective. He writes of differing skills needed for social interaction and academic success and relates this to different cognitive demands for social-interaction and academic tasks. The focus is on the task itself and context-reduced situations being used in language programmes.

For the language teacher, Cummins highlights the importance of the task itself, its learning arrangement and the demands that a lack of context could place on a less proficient learner.

The learner can be faced with a task where completing the work becomes the main concern to the extent that content becomes less important than completion of the task. If a student's language skills can not meet the task goals then there will be concern with the means of doing the task, rather than interest in the content (Frymier 1981). The rush to get the work done could mean less productive use of learning opportunities. Task designers therefore need to create tasks that involve student participation at an appropriate level. The product and the process of getting the answer could be integrated in a learning arrangement which fosters full participation with the routes to the answer.

Interaction with Tasks

Students face many kinds of work in order to succeed in a classroom. The work that they face can be analyzed by considering what has to be done in order to complete tasks in the classroom. Theoretical models and field research point to the utility of task as a unit in describing how students interact with classroom work.

Tikunoff (1985) suggests that students are faced with complex demands in order to be in what he calls a state of "functional proficiency." A summary of learning conditions that students need to master is as follows. Students need to:

1. Understand tasks knowing how to complete the product and how to correlate new information;

2. Participate productively in classroom tasks, be on task and observe the norms for the activity, while meeting teacher expectations; and

3. Obtain feedback on task completion as to its accuracy. (ibid.:19-21)

Tikunoff details the instructional demands of a task with factors that parallel Doyle's work (1979, 1983) by focusing on what a learner needs to know. Tikunoff's model has a concern with the product, an interest in conditions, and an awareness of the importance of the operations. His detailed study analyses class task demands (1985:47-51) with attention to task itself in terms of task complexity and operations, which he describes as interactional mode demands and response mode demands.

Tikunoff's analysis is important because it shows the complex requirements of even a superficially simple task and thus stresses the need for a well thought out programme of initiating learners to a new task type.

Learner-centred task analysis is central to Dale and Cuevas's (1987) study of children's mathematics difficulties and language use. They point out that their examples of children's maths difficulties need more than appraisal of the surface features of a task. The child's level of language skills and the demands placed on them by the specific domain of a maths task are at the core of analysis. The dialogue of the task quoted (1987:9-10) validates the importance of analyzing the task itself, the conditions, - primarily the functional proficiency of the child - and the operations required while the student interacts with the task.

Task as a unit of analysis is central to assessing what happens in terms of time spent in classroom learning. The teacher's expectations of what is happening may differ from student's views yet all are concerned with getting answers to set tasks. Students face many demands in completing the product and task designers need to recognise that useful classroom tasks would see the answers, or outcome as an integral part, rather than the only reason for a lesson. The product, as important as it is for learners, will only be arrived at if it matters to those concerned.

Students are concerned with the value of an answer. In fact the content of a lesson may be lost in the drive for an answer, so that task design needs to involve full participation from students, in conditions in which the process of learning is also given prominence.

Task Complexity

Task typology may need to consider task complexity as part of an analysis of learning. Information processing theory delineates how limited the processing of concepts can be. If a task has multiple components in both the procedure, the mechanics of the learning arrangements and the product it will be more difficult for learners, than a straight forward path to the answer.

Task complexity can be considered in two major ways which Segal (1982:334) describes in relation to group work.

> Two dimensions... are (1) the number of goals and (2) the number of paths to these goals (Fielder 1967, Shaw 1976)... A third dimension contributing to the complexity of the task is the amount of specific, unequivocal information given the decision maker.

This framework applies to both the procedure of a task and to the input of a task. Brown and Yule (1983) consider task complexity of listening texts and note that difficulty relates to similar informational factors. The first is the number of elements in the text and how easy or difficult it is to separate them into different concepts. The second element of difficulty is text type, which they elaborate in terms of whether elements are fixed elements or changing dynamic elements and whether the text relates to learner background knowledge. The role of information distribution and the type of information processing in describing tasks is discussed later in this paper. Complexity is also a factor in terms of the "shape" of the answer and the number of components that are involved in the expected product. Of immediate interest is the consideration of classifying language work in terms of how complex the procedures of the task are and how many parts or components the learner needs to work with.

Task analysis, based on the notion of work to be completed, may become more refined if we consider the "thinking" work needed by learners in terms of task procedural instruction, the number of components in the material and the complexity of the answer. Greater cognitive attention given to the number of components in a task will create greater linguistic work even if the language of the task appears simple. Much research needs to be done on this area of task analysis in second language teaching.

The Moderating Variables of Learning Arrangements

Task Learning and Groups

Much of what happens in a classroom will, as suggested earlier, depend on the routes to the answer. The type of learning arrangements are an important modifying influence on how routes to the answer are explored.

Several surveys of the literature (Brumfit 1984, Van Lier 1988) and assertions based on teaching experience, highlight the linguistic and pedagogical importance of group activity (Abercrombie 1970, Schmuck and Schmuck 1971, Johnson and Johnson 1975). Earlier concerns were with groups as a management device (Forrester 1968), or as an extension of essentially teacher-fronted work (Jolly and Early 1974). The important concern of how linguistically effective group learning arrangements can be has focused on linguistic analysis and the setting for task processes. Using task as a conceptual framework, Long (1989:13) sums up group work benefits:

- Group work increases the quantity of language practice opportunities.
- Group work improves the quality of student talk in several ways.
- Group activities help individualise instruction.
- Group tasks can help improve the affective climate in the classroom with the intimacy of the small group setting often being especially valuable to shy or linguistically insecure students.
- Group tasks can help motivate learners because of the advantages referred in (1) through (4).

This summary is similar to other authors' statements (Barnes 1973, Nation 1975) and describes positive aspects that are widely known in the SLA field. However group work arrangements are only as effective as the arrangements of information that align with the group arrangement and set up useful processing of linguistic information.

Limitations in Small Group Work

Some detailed studies suggest limitations with the implementation of small groups. Gerleman (1987) notes that small groups may not fully achieve their aims. Her observational study of fourth grade mathematics classes found that group activities did not necessarily create equal content learning opportunities for all students. Secondly, control of student behaviour became a problem, with teachers' attention being spread too widely. Management problems would be fewer in pair work where the grouping retains the features of cooperation within interaction with full student participation, or in group work with systems of individual accountability (Hall and Jacobs, in press).

Participation and performance concern Slavin (1980) in his study of cooperative learning. An analysis of reward structures for a varied range of cooperative techniques shows that the larger the group, the greater the reduction in the degree to which individual performance relates to individual achievement. A group member can cover for lesser effort from another student doing the group activity. Group members may pull each other up with appropriate group rewards, but individual performances can vary in effort and learning.

Slavin notes that group member praise only works on certain tasks, suggesting that group techniques are effective when they focus on structured tasks, high student autonomy and defined individual accountability. This finding links with

Doyle's (1983) assertion that accountability drives the task system. It also clarifies an assertion that is often heard in school staff rooms; "He only does enough work to get by." This would suggest that how the learner perceives accountability, whether in group work or in individual tasks, is an essential part of task design. One comes close to an ESP approach in aligning the settings for language acquisition with learner awareness of the measurement of successful task completion. This may be accountability within the classroom task, or in terms of a target task outside the classroom, but even the best groups and engaging tasks will only work for the majority by building accountability and feedback into the task design.

Task and Operations

Background

Analysis of school work in recent years has begun to relate to operations generated by the work or task. Research on the work of school tasks has overlapped with psychology's concern with cognitive processes (Anderson, Spiro and Montague 1977, Calfee 1981, Chamot and O'Malley 1987). In many schools this same interest in operations correlating with tasks has led to widespread interest in "process writing " and lead to a concern with the processes of learning that tasks can engender. The aptly named "process" approach is currently widely applied and researched (Graves 1975, 1983, O'Rourke and Phillips 1989). There is much practical and theoretical interest in the operations or processes generated by different types of tasks.

Task types and their influence in the process of learning with different kinds of class work has been a feature of some areas of SLA research. Task types may be considered as a condition of learning yet much of the research has focussed on the operations that various types of task engender. Descriptions based on classroom research into the operations or processes may aid the classroom teacher in designing effective tasks.

A Framework for Analysis of Process

Doyle (1983) focuses on process with a general framework for academic tasks by stating that "Academic tasks embedded in the curriculum can be differentiated in terms of general categories of cognitive operations that are involved in task accomplishment" (1983:162-3). His classification for general academic analysis is relevant to language teaching.

Four types of tasks are described:

1. Memory Tasks are activities in which the student recognises or reproduces previously known items.

2. Procedural or Routine Tasks are activities in which the student applies formulae or algorithms.

3. Comprehension Tasks are activities involving transformation, new procedures or inference.

4. Opinion Tasks are activities in which the student states preferences or values. (1983:163-173)

Doyle's analysis of product, givens and operations with the above types of task classification has contributed to a growing field of research where the concern is learner centred task-based analysis of what the student is actually doing. (Duff 1986, Mergendoller, Marchman, Mitman and Packer 1988, Anderson, Stevens, Prawat, Nickerson 1988).

Analyzing task, student and teacher operations has occurred in a wide range of academic areas. Crookes (1986) presents a detailed summary of task characteristics in relation to second language classrooms. In general classrooms, with or without ESL learners, the interplay of task forms and performance is just as observable. An example of task forms influencing performance is discussed in recent American elementary school research. Blumenfeld and Meece (1988) used four science lessons with 194 fourth through to sixth grade students. The lessons varied in the level of cognitive content, procedural complexity and learner arrangements. Using questionnaires, they measured student task involvement, teacher influence and the use of cognitive strategies. They found that cognitive engagement which they defined as "self regulated learning such as attention, connecting (and) planning" (ibid.:239), involved more learning strategies in tasks of high cognitive difficulty.

The importance of form was seen, in that procedurally complex tasks involved lower cognitive engagement. Small group work involved less engagement as well. Students' interview responses suggest that what creates interest or involvement, is related to the procedures or forms of a task, such as graphing or drawing (ibid.:246). This raises interesting questions in both the analysis and the productive use of tasks in relation to content. The researchers state that student concern focused on what had to be done, rather than on the content. The outcome itself is seen as being of primary importance in how students perceive classroom work.

The Influence of Task Types on Learner Operations

Discourse Types

Interest in task types and changes in interlanguage in pair work is central to Crookes and Rulon's (1985) study of native speaker (NS) and non-native speaker (NNS) pairs. The central issue was the quality and quantity of NS"feedback" to NNS. Two "problem solving" tasks were compared with "free conversation." The first task involved partners agreeing on which, out of four items, was "The Odd Man Out." The second involved visual discrimination with a "Spot the Difference" task.

Modification of NNS language by NS feedback was more frequent in problem solving tasks than in free conversation. It was hypothesised, in a variation of the research questions on time on task, that the repetition of vocabulary items and topic areas could be a major factor. In other words, repetition of the same linguistic items may be more important than the nature of discourse generated by a task type. In "Spot the Difference", there were long stretches of topic centred discourse with repetition, but the activity had less effect on interlanguage exchange than the other problem solving task. Designing pair work or group interaction tasks chiefly on the basis of the quantity of repetitions would not seem fruitful.

Duff (1986) analyzed NNS dyads using "problem solving" and "discussion" tasks. She argues for the analysis of the task itself into discourse that centres on "convergent discourse" and "divergent discourse." In convergent discourse, learners have a shared goal in that they are cooperating in processing the same product. Tasks such as free discussion in pairs must, by their very nature, set up different viewpoints for each student. The task would then have "opposite or independent goals for each member" and be "divergent" (ibid.:150). Chaudron (1988:109) describes Duff's findings by analyzing the research as that with "task type as an independent variable."

> Calculating ratios of behaviours to total turns, she found the problem-solving task significantly superior to debates only on (1) the rate of questions posed by the subjects (2) the rate of "referential" questions and the role of confirmation checks.

Duff's results show that convergent tasks lead to more exchanges, more questioning, and a range of checking procedures. It is the task type which creates opportunities for questioning and checking, an aspect of negotiation which may be significant. The length and structure of the exchanges in divergent tasks means fewer possibilities for negotiation, in the form of checks and questions.

These findings are confirmed by Tong-Frederick's (1984) experimental analysis of three different kinds of oral communication. She compared "a goal directed problem solving" activity, a role play,and "authentic/natural" interaction (1984:133) using six pairs of students. The students did all activities. It was hypothesised that the lexical complexity of both a role play and a naturalistic activity of finding out what another student had done on the previous day, would lead to a wide range of forms and vocabulary items. Lexis would be limited in the problem solving task as a specific, rather than as a wide-ranging problem. Experimental data contradicted this.

> In the problem solving activity, there were many instances of students rephrasing, defining the task to their partner, drawing inferences, checking, disagreeing, agreeing, organising, contradicting and evaluating. It was the task type itself which was influential (ibid.:142).

Furthermore students were interviewed and stated a clear preference for this form of pair work learning in that a cognitive challenge was motivating.

Open Tasks and Closed Tasks

Using the product of a task as the criterion, Long (1989) argues for a fundamental division of pedagogic tasks into "open" and "closed" tasks. Open tasks are those in which participants know there is no predetermined correct solution, but instead, a wide (in some cases, infinite) range of acceptable solutions" (1989:24). Open tasks involve consensus and produce "divergent discourse" in Duff's terms.

> A 'closed task' requires that speakers attempt to reach either a single correct solution or one of a small finite set of correct solutions determined before hand by the designer of the task and again (crucially) known to the participants to have been so determined" (ibid.:25).

Long states that closed tasks foster more negotiation as students know that finishing the task involves finding a specific product or answer. Students are extended to find the specific right answer. The answer or product is therefore central to two-way tasks being a principled use of the student perception of classwork, as described when analyzing task and the important role of the product.

Two-Way Tasks Compared to One-Way Tasks

The dyad is a face to face setting where information can be exchanged. Tasks can be examined in the light of how ideas are exchanged. Long (1980) defines a "two-way task" as one where students hold different information which must be shared to complete the activity. This technique has been given various names - dycoms (Byers 1973), jigsaw groups (Aronson et al. 1975), two-way information gap tasks (Doughty and Pica 1986), and split information technique (Nation and Thomas 1988, Hall 1992).

> "Two-way" tasks involve exchanges of information in which both participants have unique information which must be shared to process the task (Gass and Varonis 1985, Nation and Thomas 1988) in contrast to "one-way" tasks. Often in one-way tasks, one learner or the teacher holds all the information and has to convey it. Roles are defined by the nature of one-way tasks as passive and active users of language at one extreme or at the

best, a situation of exchange and turn imbalance. Much teacher-fronted teaching involves tasks of this type. Two-way tasks produce interactional features that differ from one-way tasks. A critical consideration in analyzing tasks in this way is the influence that the information distribution has on interaction and talk which could link to language acquisition.

Long (1980) analyzed the amount of interactional changes in both one-way and two-way tasks by analyzing discourse in dyads. He compared the discourse of 16 NS-NSS dyads with NS-NS dyads. Interactional features in the form of checks, repetitions and requests occurred in both one-way and two-way tasks. There were, however, greater differences between NS-NS and NS-NSS conversation in the two-way tasks. It could be argued that the native speakers, working with a need to process a specified product, had to modify their language to complete the task. In other words, the holding of specific unique information and a convergent discourse structure (Duff 1986) may have led to the NS modification and negotiation.

Pica and Doughty (1985a) support Long's findings with an analysis of teacher-fronted discussions compared with discussions of small groups. Each group was of four learners. The tasks were values based activities which are not uncommon in communicative methodology, (Brumfit 1980, Ellis 1984b). Tasks involved reaching a consensus. In the small group the students had to decide which of the designated twenty-first century families could best adopt a child. Learners shared the same information and the researchers suggested that as students did not hold unique information, they did not have to participate equally or make all their language clear.

This is a finding which many classroom teachers would agree with and it may account for the reluctance of some teachers to use small groups. Johnson and Johnson (1985), describe the lack of use of small groups, noting the management and involvement problems which may have more to do with task types than with classroom management. A personal observation is that a shy student will not take part in group processes if they "don't really need to do all the work."

In a later study more directly concerned with the role of one-way and two-way information gap tasks, Pica and Doughty (1986), returned results which highlight the importance of each learner holding unique information. Task and learning arrangement components mark the study as a development of earlier work (1985a, 1985b). Firstly, all tasks had a description of information exchange, either optional in the teacher mode or required. Secondly, the major comparison of teacher-fronted versus group work was supplemented by a third pattern, the student dyad. The researchers showed a direct interest in dyads. Greater modification of language occurred in the "required information" tasks where unique information was held by each learner, whether it was teacher-fronted or part of group work. One notes that two-way tasks provide many opportunities for student to student interaction. The dyadic two-way task could lead to students having to negotiate language and content in a meaningful way (Hall 1992). What is of interest is the importance of the roles that are created by learning arrangements and the importance of information distribution in tasks.

Critical factors in classifying a range of effective tasks may be the type of task as determined by the discourse generated, the learning arrangement, the information distribution and the roles of learners that the task sets up. The practicalities of applying "task" as a unit for the practice and analysis of teaching are addressed in the following section.

Applying Task As A Unit of Teaching Analysis

Task may be practically applied as a unit of teaching analysis by relating the reasons for a classroom to the learners and their needs. The learners' thinking and the preoccupation that the drive for the answer creates can be incorporated into a framework of task as a unit of teaching analysis.

TASK-BASED TEACHING CYCLE

Target Task

Role of Learners and Teachers

Input/Identify Skills

Pedagogic task

Product

Evaluation

Target Task

The widely used ESP concept of target task recognises that learners' motivation can be both integrative and instrumental in that the learner may often be meeting needs and answering demands that stem from outside the classroom. The relevance of an EFL programme will often be seen by learners in terms of "Where", "When" "With Whom" and "How" English will be applied.

Target task choice will require analysis of communicative situations, purposes and functions of relevance to the majority of learners in a class.

Roles

Roles of the learners and teachers will determine much of the nature of learning. As has been suggested, different learning arrangements will stimulate different processes of language acquisition. The roles of the learners in pair work are different from those in teacher-fronted teaching and the role of a teacher in leading intensive reading is different from that in group-based fluency work. How language work proceeds depends on the arrangement for performing the task so that roles or learning arrangements are important.

Input/Identify Skills

Input which is appropriate and motivating is critical as the materials need to pull together both target needs, learner interest and the final product. In choosing input one needs to identify skills that can function within the material. The material obviously needs to be appropriate, intellectually challenging to learners and reflective of meeting learners needs in terms of outside the classroom and to meet the challenge of the pedagogic task.

Pedagogic Task

Given that products or answers are important, much of the answering along the route to the product is the real learning. The process is mediated by learner motivation, the interest in the materials and the cognitive processes that are engaged in. The task can create depth of processing, be presented with clear procedures and echo the target tasks that learners face. Choosing a task may involve an analysis of the information arrangement and the thinking that can occur in different formats, as well as well known features of well chosen vocabulary and topics.

Product

The product or answer is an important concern. While there has been much focusing on the process of learning one can not deny the drive to reach an answer. A product or answer may be seen as valid if it is working towards an authentic form or is applicable in that it appropriates target task completion as well as a sense of achieving the piece of classroom work. Comprehensible output and productive use of language are important considerations in teaching choices.

Evaluation

Evaluation by both learners and course providers can provide insight into the relationship between meeting learners needs through useful learning arrangements, motivating input, useful types of tasks and productive learning.

To summarise, one can evaluate the classroom's relevance through meeting needs and aligning target tasks with that which makes up a lesson unit. The lesson unit or pedagogic task will be motivating if students see its relevance. The relevance is not only in the "what" of learning, the input, but also in the "how" - the learning arrangement being a framework for processing language, ideas, skills and text. The processes of the task need to be aligned with the product, so that the answer will reflect the amount of thinking and language generated by processes. Our evaluation as part of the cycle will ensure that the classroom is task based, learner oriented and productive in aiding the learning and acquisition of language. In structuring and organising the drive for answers we work with tasks, and hopefully choose and facilitate processes that are rewarding for learners in a changing world.

References

Abercrombie, M.L.J. (1970). *Aims and Techniques in Group Teaching.* London: Society for Research into Higher Education.

Anderson, J.R. (1980). *Cognitive Psychology and its Implications.* San Francisco: W.H. Freeman.

Anderson, L. (1981). Short-Term Student Responses to Classroom Instruction. *Elementary School Journal* 82:97-108.

Anderson, L.M., Stevens, D.D., Prawat, R.S., & Nickerson, J. (1988). Classroom Task Environment and Students Task-related Beliefs. *Elementary School Journal* 88(3):281-295.

Anderson, R.C., Spiro, R.J., & Montague, W.E. (1977). *Schooling and the Acquisition of Knowledge.* Hillsdale NJ: Lawrence Erlbaum Associates.

Aronson, E. et al. (1975). The Jigsaw Route to Learning and Liking. *Psychology Today* 8(9):43-50.

Barnes, D. (1973). *Language in the Classroom.* Bletchley, England: Open University Press.

Bennett, N. (1988). The Effective Primary School Teacher: The Search for a Theory of Pedagogy. *Teaching and Teacher Education* 4(1):19-30.

Bennett, S.N. (1982). Time to Teach: Teaching Learning Processes in Primary Schools. *Aspects of Education* 27:51-70.

Berliner, D.C. (1987). Knowledge is Power: A Talk to Teachers About a Revolution in the Teaching Profession. In D.C. Berliner & B.V. Rosenshine (eds.). *Talks to Teachers.* New York: Random House.

Bloom, B.S. (1976). *Human Characteristics and School Learning.* New York: McGraw-Hall.

Blumenfield, P.C. & Meece, J.L. (1988). Task Factors, Teacher Behaviour and Students Involvement and Use of Learning Strategies in Science. *The Elementary School Journal* 88(3):235-250.

Blumenfield, P.C., Mergendoller, J.R., & Swarthout, D.W. (1987). Task as a Heuristic for Understanding Student Learning and Motivation. *Journal of Curriculum Studies* 19(2):135-148.

Breen, M.P. (1987). Learner contributions to task design. In C.N. Candlin & D. Murphy (eds.), *Lancaster Practical Papers in English Language Education: vol 7. Language Learning Tasks* (pp. 23-46). Englewood Cliffs, NJ: Prentice Hall.

Brophy, J. & Good, T. (1986). Teacher Behavior and Student Achievement. In M. Wittrock (ed.), *Handbook on Research on Teaching* (pp. 328-375). New York: MacMillan.

Brophy, J., Rohrkemper, M., Rashid, H., & Goldberger, M. (1983). Relationships Between Teachers' Presentations of Classroom Tasks and Students' Engagement in Those Tasks. *Journal of Educational Psychology*, 75:544-552.

Brumfit, C.J. (1984). *Communicative Methodology in Language Teaching.* Cambridge: Cambridge University Press.

Brumfit, C.J. (1980). *Problems and Principles in English Teaching.* Oxford: Pergamon.

Burns, R.B. (1984). How Time is Used in Elementary Schools: the Activity Structure of Classrooms. In L.W.Anderson (ed.), *Time and School Learning* (pp. 91-127). New York: St. Martins Press.

Byers, B.H. (1973). Testing Proficiency in Interpersonal Communication. *RELC Journal*, 4(2):39-47.

Calfee, R. (1981). Cognitive Psychology and Educational Practice. In D.C. Berliner (ed.), *Review of Research in Education*, Vol.9. Washington D.C.: American Educational Research Association.

Candlin, C.N. (1987). Towards task-based language learning. In C.N. Candlin & D. Murphy (eds.), *Lancaster Practical Papers in English Language Education, Vol. 7. Language learning tasks*, (pp. 5-22). Englewood Cliffs, NJ: Prentice Hall.

Carroll, J.B. (1963). A Model of School Learning. *Teachers College Record*, 64:723-733.

Carroll, J.B. (1985). The Model of School Learning: Progress of an Idea. In C.W. Fisher & D.C. Berliner (eds.), *Perspectives on Instructional Time*, (pp. 29-58). New York: Longman.

Carter, K. & Doyle, W. (1982). Variations in Academic Tasks High and Average Ability Classes. Presented at the Annual Meeting of the American Education Research Association. New York: March 1982.

Chamot, A.V. & O'Malley, J.M. (1986). *A Cognitive Academic Language Learning Approach : An ESL Content-Based Curriculum.* Rosslyn, VA: National Clearinghouse for Bilingual Education.

Chaudron, C. (1988). *Second Language Classrooms: Research on Teaching and Learning.* Cambridge: Cambridge University Press.

Clark, C.M., & Yinger, R.J. (1979). Teachers Thinking. In P.L.Peterson & H.J.Walberg (eds.), *Research on Teaching.* Berkeley, CA: McCutchan.

Croll, P. (1988). Teaching Methods and Time on Task in Junior Classrooms. *Educational Research*, 30(2):90-98.

Crookes, G. (1986). *Task Classification: A Cross Disciplinary Review.* Technical Report No 4: Center for Second Language Classroom Research. Hawaii: University of Hawaii at Manoa.

Crookes, G & Rulon, K.A. (1985). *Incorporation of Corrective Feedback in Native Speaker/Non-Native Speaker Conversation.* Technical Report No 3: Center for Second Language Classroom Research. Manoa, Hawaii: Social Science Research Institute.

Cummins, J. (1984). *Bilingualism and Special Education: Issues in Assessment and Pedagogy.* San Diego: College Hill Press.

Dale, T.C. & Cuevas, G.J. (1987). Integrating Language and Mathematics Learning. In J. Crandall (ed.), *ESL Through Content Area Instruction*, (pp. 9-54). Prentice Hall Regents: New Jersey.

Denham, G. & Lieberman, A. (eds.) (1980). *Time To Learn.* Washington: National Institute of Education.

Doughty, C. & Pica, T. (1986). Information Gap Tasks: Do they Facilitate Second Language Acquisition? *TESOL Quarterly* 20(2):305-325.

Doyle, W. (1983). Academic Work. *Review of Educational Research* 53(2):159-199.

Doyle, W. (1986). Classroom Organisation and Management. In M. Wittrock (ed.), *Handbook of Research on Teaching,* (pp. 392-431). New York: Macmillan.

Duff, P.A. (1986). Another Look at Interlanguage Talk: Taking Task to Task. In R.R Day (ed.). *Talking to Learn: Conversation in Second Language Acquisition,* (pp. 147-181). Rowley, MA: Newbury House.

Ellis, R. (1984). Communication Strategies and the Evaluation of Communicative Performance. *ELTJ* 38(1):39-44.

Fielder, F.E. (1967). *A Theory of Leadership Effectiveness.* New York: McGraw-Hill.

Forrester, J. (1968). *Teaching Without Lecturing.* Oxford: University Press.

Frymier, J. (1981). Learning Takes More Than Time on Task. *Educational Leadership,* 38:634-649.

Gaies, S.J. (1983). The Investigation of Language Classroom Processes. *TESOL Quarterly* 17(2):205-217.

Gass, S. & Varonis, E.M. (1985). Task Variation and Non-native/Native Negotiation of Meaning. In S. Gass & C. Madden (eds.). *Input in Second Language Acquisition* (pp. 149-161). Rowley, MA: Newbury House.

Gerleman, S.L. (1987). An Observational Study of Small Group Instruction in Fourth Grade Mathematics Classrooms. *Elementary School Journal,* 88(1):3-28.

Gettinger, M. (1984). Individual Differences in Time Needed for Learning: A Review of Literature. *Educational Psychologist,* 19:15-29.

Gettinger, M. (1989). Effects of Maximising Time Spent and Minimizing Time Needed for Learning on Pupil Achievement. *American Educational Research Journal,* 26(1):73-91.

Graves, D.H. (1975). An Examination of the Writing Processes of Seven Year Old Children. *Research in the Teaching of English,* 9:227-241.

Graves, D.H. (1983). *Teachers and Children at Work.* New Hampshire: Heinemann.

Hall, S.J. (1992). Using Split Information Tasks to Learn Mathematics Vocabulary. *Guidelines*. 14(1):72-78.

Harley, B., Allen P., Cummins J. & Swain M. (eds.), (1990). *The Development of Second Language Proficiency*. Cambridge: Cambridge University Press.

Hyland, K. & Hyland F. (1992). Go for Gold: Integrating Process and Product in ESP. *English for Specific Purposes*. 225-242.

Johnson, R.T., & Johnson, D.W. (1985). Student-Student Interaction: Ignored but Powerful. *Journal of Teacher Education* 36(4):22-26.

Jolly, D. & Early, P. (1974). *Group Work in English Language Teaching*. The British Council: Belgrade.

Littlejohn, A.P. (1983). Increasing Learner Involvement in Course Management. *TESOL Quarterly* 17(4):595-608.

Long, M.H. (1980). Inside the 'Black Box': Methodological Issues in Classroom Research on Language Learning. In H.W. Seliger & M.H. Long (eds.), *Classroom Oriented Research in Second Language Acquisition*. Rowley, MA: Newbury House.

Long, M.H. (1985). A Role for Instruction in Second Language Acquisition: Task Based Language Teaching. In Hyltenstam, K. & Pienemann, M. (eds.), *Modelling and Assessing Second Language Acquisition* (pp. 77-99). San Diego: College-Hill Press.

Long, M.H. (1989). Task, Group and Task Group Interactions. Plenary Address to RELC, Regional Seminar. April 1989, Singapore.

Marx, R.W., & Walsh, J. (1988). Learning from Academic Tasks. *Elementary School Journal*, 88(3):207-219.

Mergendollar, J.R., Marchman, U.A., Mitman, A.L., & Packer, M.J. (1988). Task Demands and Accountability in Middle Grade Science Classes. *The Elementary School Journal*, 88(3):251-265.

Mohan, B.A. (1979). Relating Language Teaching and Content Teaching. *TESOL Quarterly*, 13(2):171-182.

Mohan, B.A. (1986). *Language and Content*. Reading, MA: Addison Wesley.

Murphy, J.M. (1991). Oral Communication in TESOL: Integrating Speaking, Listening and Pronunciation. *TESOL Quarterly*, 25(1):51-75.

Nation, I.S.P. (1975). *Language Teaching Techniques*. Wellington: English Language Institute, Victoria University.

Nation, I.S.P. (1983). *Teaching and Learning Vocabulary*. Wellington: English Language Institute, Victoria University.

Nation, I.S.P. & Thomas, G.I. (1988). *Communication Activities*. Wellington: English Language Institute, Victoria University.

Nunan, D. (1991). Communicative Tasks and the Language Curriculum. *TESOL Quarterly*, 25(2):279-295.

O'Rourke, A. & Philips, D. (1989). *Responding Effectively to Pupils' Writing*. Wellington: New Zealand Council For Educational Research.

Peterson, P.L., Marx, R.W., & Clark, C.M. (1978). Teacher Planning, Teacher Behavior and Student Achievement. *American Educational Research Journal*, 15:417-432.

Pica, T. & Doughty, C. (1985a). Input and Interaction in the Communicative Language Classroom: A Comparison of Teacher Fronted and Group Activities. In S.M. Gass & C.G. Madden (eds.), *Input in Second Language Acquisition*, (pp. 115-132). Rowley, MA: Newbury House.

Pica, T. & Doughty, C. (1985b). The Role of Group Work in Classroom Second Language Acquisition. *SSLA* 7:233-248.

Pica, T., Doughty C. & Young R. 1986. Making Input Comprehensible: Do Interactional Modifications Help? *ITL Review of Applied Linguistics*, 72:1-25.

Pienemann, M. & Johnston, M. (1987). Factors influencing the development of language proficiency. In D. Nunan (ed.), *Applying second language acquisition research*, (pp. 45-141). Adelaide, Australia: National Curriculum Resource Centre.

Prabhu, N.S. (1987). *Second Language Pedagogy*. Oxford: Oxford University Press.

Richards, J.C. & Rodgers T.S. (1986). *Approaches and Methods in Language Teaching*. Cambridge: Cambridge University Press.

Richards, J.C. (1990). *The Language Teaching Matrix*. Cambridge: Cambridge University Press.

Schmuck, R.A. & Schmuck, P.A. (1971). *Group Processes in the Classroom*. Dubuque, IO: W.C. Brown.

Segal, V.A. (1982). The Cyclical Nature of Decision Making. *Small Group Behavior*, 13(3):333-348.

Shavelson, R.J., & Stern, P. (1981). Research on Teachers Pedagogical Thoughts, Judgements and Behavior. *Review of Educational Research*, 51(4):455-498.

Shaw, M.E. (1976). *Small Group Dynamics: The Psychology of Small Group Behavior*. New York: McGraw-Hill.

Slavin, R. (1977). A Student Team Approach to Teaching Adolescents with Special Emotional and Behavioral Needs. *Psychology in the Schools*, 14:77-84.

Smith, E.L. & Sendleback, J. (1979). Teacher Intentions for Science Instruction and Their Antecedents in Program Materials. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.

Stahl, S.A. & Vancil, S.J. (1986). Discussion is What Makes Semantics Maps Work in Vocabulary Instruction. *The Reading Teacher*, 40(1):62-67.

Swaffar, J.K., Arens, K. & Morgan, M. (1982). Teacher Classroom Practices: Redefining Method as Task Hierarchy. *Modern Language Journal*, 66(1):24-33.

Tikunoff, W.J. (1985). *Developing Student Functional Proficiency: A Teachers Casebook*. Gainesville, FL: University of Florida.

Tong-Fredericks, C. (1984). Types of Oral Communication Activities and the Language They Generate: A Comparison. *System*, 12(2):133-146.

Ushimaru, A. (1992). EFL Learners Talking to Each Other: Task Types and the Quality of Output. *JALT Journal*, 14:1.

Van Lier, L. (1988). *The Classroom and the Language Learner*. New York: Longman.

Wang, M.C. (1987). Individual Differences and Effective Schooling. *Professional School Psychology*, 2:53-66.

White, L. (1987). Against comprehensible input: The input hypothesis and the development of L2 competence. *Applied Linguistics*, 8(1):95-110.

Young, R. (1988). Variation and the interlanguage hypothesis. *Studies in Second Language Acquisition*, 10(3):281-302.

Zahorik, J.A. (1975). Teachers' Planning Models. *Educational Leadership*, 33:134-139.

Zobl, H. (1985). Grammars in search of input and intake. In S.M. Gass & C.G. Madden (eds.), *Input in second language acquisition* (pp. 329-344). Rowley, MA: Newbury House.

USING A FUNCTIONAL APPROACH IN ASSESSING WRITTEN TEXTS

David Nunan

Introduction

This paper presents a case for the use of a functional approach to the analysis and assessment of written language in schools. A key concept in the analytical scheme described in the paper is that of 'genre'. "The term 'genre' has been used for many years to refer to different styles of literary discourse ... and highlights the fact that different types of discourse can be identified by their overall 'shape' or 'generic structure'. In recent times, the term has been adapted by functional linguists to refer to different types of communicative events." (Nunan 1993) It is this more recent adoption which is taken up in this paper.

In the paper I shall provide a brief outline of a functional linguistic model. I shall then describe two curriculum projects with which I have been involved. The first, an evaluation of the Disadvantaged Schools Writing Project, was carried out in Sydney. The second, the South Australian Needs Assessment Project (SNAP), took place in Adelaide. Despite their different orientations, both projects shared a common ideological base in that they were both based on a functional view of language derived from systemic-functional linguistics. The paper concludes with a discussion of the advantages and disadvantages of using a functional approach to the assessment of written language.

Functional Linguistic Models and the Notion of 'Genre'

Functional models of linguistic analysis are developed to account for relationships between the forms of the language, and the various uses to which the language is put. The systematic relationship between language structure and function is described by Halliday (1985) in the following way:

> Every text - that is, everything that is said or written - unfolds in some context of use; furthermore, it is the uses of language that, over tens of thousands of generations, have shaped the system. Language has evolved to satisfy human needs; and the way it is organised is functional with respect to those needs - it is not arbitrary. A functional grammar is essentially a 'natural' grammar, in the sense that everything in it can be explained, ultimately, by reference to how language is used.

A key concept for many working within this functional perspective is 'genre'. A 'Genre' is a particular type of oral or written communication such as a narrative, a casual conversation, a poem, a recipe or a description. Different genres are typified by a particular structure and by grammatical forms that reflect the communicative purpose of the genre in questions.

The 'genre' theorists, then, argue that language exists to fulfil certain functions, and that these functions will determine, not only the grammatical items which appear in a text, but also the overall shape or structure of the discourse which emerges as people communicate with one another. In other words, it will have certain predictable stages. The communicative purpose will also be reflected in the basic building blocks of the discourse, that is, the words and grammatical structures themselves. In other words, different types of communicative events result in different types of discourse, and these will have their own distinctive characteristics. Some events result in sermons, others in political speeches, and yet others in casual conversations. While each sermon, political speech and casual conversation will be different, each discourse type will share certain characteristics which will set it apart from other discourse types.

I should like to illustrate this with reference to two authentic texts, one spoken and one written. (The examples are taken from Nunan 1993:50-51)

1. Written text:

   Hanging bungle uncovered
   By Geoff Easdown, Mike Edmonds and Barry MacFadyn

   MELBOURNE: A sensational development in the case of Ronald Ryan, the last man hanged in Australia, shows a bungle almost certainly cost him his life.

   It was revealed last night that four letters written by jurors in the trial, appealing for Ryan not to be hanged, were never sent to the Victorian Cabinet which decided to execute him.

   And a member of the Victorian Cabinet that voted 11-4 to hang Ryan, Sir Rupert Hamer, says the mercy pleas by four jury members could have saved Ryan. (The Advertiser January 6, 1992)

2. Spoken text

   A: What did you do last night?

   B: Well, Mum and Dad went out so we went to Marg's to sleep, and Sarah wouldn't go to sleep, and she wanted to ring Mum, and Marg said she couldn't, and so she cried, and so Marg combed her hair, and then she went to sleep. She was really naughty ....

   A: What time did she go to sleep?

   B: mmm - 'bout one o'clock. (Author's data)

The written text is an extract from a leading article in a newspaper. Its generic structure (at least the structure of the extract) is as follows:

Title
Author (s)
Location
Argument
Supporting detail
Supporting detail

Linguistically, the piece contains an agentless passive 'It was revealed ...' It also contains emotively charged words such as 'sensation', 'bungle', 'mercy', 'plea'. In terms of its layout and physical appearance, the text contains a large, eye-catching headline. The columns and assignment of each sentence to a separate paragraph are designed to make the piece easy to read.

The spoken extract is taken from a conversation between a girl and her grandmother, and contains a recount. According to functional linguists, recounts consist of a sequence of events which are initiated by an introduction and orientation, and which end with a comment and conclusion. We can see that, with the exception of a conclusion, this recount conforms to the proposed generic structure. Grammatically, recounts are characterised by the simple past tense, and the use of specific reference to people and places. We can see that this recount, in addition to its generic structure, also contains the grammatical items of simple past tense and specific reference.

Introduction:	Well, Mum and Dad went out
Orientation:	so we went to Marg's to sleep,
Event:	and Sarah wouldn't go to sleep,
Event:	and she wanted to ring Mum,
Event:	and Marg said she couldn't,
Event:	and so she cried,
Event:	and so Marg combed her hair,
Event:	and then she went to sleep.
Comment:	She was really naughty.

At present, linguists are studying different text and discourse types in an effort to identify their underlying generic structure, and the linguistic elements which characterise them. In addition to identifying generic structure, and linguistic features, genre analysts also look at other discourse features such as topicalisation, the use of reference, and the operation of given/new structures in text.

What are some of the practical applications of this model? In the next two sections, I shall outline two recent curriculum innovations which have employed the analytical tools offered by genre theory and functional grammar.

Case Study 1: An Evaluation of the Disadvantaged Schools Project Writing Package

The first case I wish to cite is an evaluation of an innovative curriculum program focusing on the development of written language in primary and secondary schools (although we shall only look at the primary school data here.) I was commissioned, along with three colleagues at the National Centre for English Language Teaching and Research, Macquarrie University, to evaluate this innovation. (For a detailed description of the project, see Nunan 1992.)

The innovation had six objectives, as follows:

1. That students demonstrate improvements in their ability to respond effectively to the writing demands of the curriculum. Specifically, they will be able to perform effectively in written class assignments, make effective notes, do independent research, complete written homework assignments, participate fully in classroom discussions about writing.
2. That students understand the criteria by which their writing is being assessed and act on their writing to meet these criteria.
3. That there is evidence to demonstrate the positive impact of the teaching/learning cycle, known as the 'curriculum genre', on students' verbal and reading abilities.
4. That teachers participating in the genre writing package be able to identify examples of the following genres: recount, report, procedure, explanation, exposition, discussion and narrative.
5. That teachers will be aware of the significant language features of the genres listed above.
6. That teachers will be able to apply their knowledge of genre theory to identify the schematic structures and significant language features of genres other than those identified above.

The evaluation had three overall purposes:

1. To assess the impact of the 'package' on children's writing.
2. To evaluate the impact of the package on teachers'
   a) capability to assess the effectiveness of students' writing
   b) pedagogy
   c) knowledge of the social functions of language
3. To identify which elements of the package have been most beneficial and which require amendment.

While the evaluation focused on the impact of the package on the practices of the teachers, as well as on the children's writing, I shall confine myself here to the learners who took part in the project.

Data collection methods:

1. A detailed questionnaire completed by teachers involved in the project;

2. Focussed interviews with teachers and other key personnel;

3. Observation and analysis of lessons;

4. Analysis of samples of students' writing from schools taking part in the innovation and also from a sample of the schools not involved in the innovation.

Timeline:

MAY: Appoint principal researcher. Devise questionnaire and distribute to schools

JUNE: Interview consultants and authors of inservice package on goals, nature and implementation of package. Conduct literature review. Identify 'non'-package schools to act as control. Collect samples of students' writing from package and non-package schools.

JULY: Collate responses to questionnaire and select schools for further evaluation. Structured interviews with teachers, students and parents. Recording of sample lessons and collection of written texts relating to these lessons

AUG: Complete interviews, recordings of sample lessons and collection of texts. Begin data analysis and evaluation of students' writing

SEPT: Complete data analysis and evaluation of students' writing. Begin drafting report.

OCT: Submit draft report.

NOV: Revise report. Submission of final draft.

When the data were in, we discovered that we had around 1,500 pieces of children's writing from the package schools, and a number of comparison schools. We had two weeks in which to analyse and evaluate the samples. The first task was to develop assessment criteria based on the linguistic model underlying the innovation. A complete analysis was impossible, so we confined ourselves to three characteristic features: schematic structure, topic development, and reference. These are described below.

Assessment criteria:

1. Schematic structure: Is the schematic structure appropriate for the genre of the text? (In the model, it is argued that texts written for different purposes will exhibit different patterns of overall organisation and text structure.)

2. Topic development: Does the writer explicitly identify the topic, and was the topic developed appropriately? (If the writer fails to develop the text topic or switches from one topic to another, then the text is confusing and difficult to follow.)

3 Reference: Does the writer use reference appropriately? (Appropriate use of reference is an indicator of text cohesion and an indicator that the writer has a sense of the 'decontextualised' nature of writing in comparison to speaking.) (Nunan, 1992:204)

Schematic structure of successful text:

Structure	**Clauses**	**Text: The Skull and the Skeleton**
Orientation	1	One day there was a poor orphan girl
	2	She had to work with her stepmother
	3	Her hands were going to skin and bones
	4	So she decided to run away
	5	She saw a castle
	6	So she knocked on the door tap tap tap
	7	A skull with no body opened the door
	8	and he said "yes"
Complication	9	The girl told the skull [what had happened to her]
	10	She stepped into his castle
	11	She saw a body without a skull
	12	She knew that it belonged to the skull
	13	And the skull told the girl [what had happened]
	14	The (sic) they had dinner
	15	She stayed two night [sic]
Resolution	16	and she kissed the skull
	17	They got married
	18	They lived happily after

Commentary on text:

[This text] like many other narratives collected for this analysis contains no evaluation and reveals that this young writer, like many others, lacks full control of the narrative genre. However, for the purpose of this evaluation of young children's writing the essential stages of the narrative have been taken to be orientation, complication and resolution, and hence is assessed as satisfying criterion 1. The topic of [the text] is developed in the sense that the adventures of the 'poor orphan girl' are related to the meeting of the 'skull' and subsequent finding of its disengaged body. While the logical sequence of some events in the Narrative such as the skull telling 'what had happened' and 'having dinner' are not especially clear, there is enough information about the skull and skeleton for the reader to follow both the sequencing of events and the connection between complication and resolution. Hence the text is considered successful in terms of criterion 2. Reference is used appropriately in the text. The major participants are explicitly introduced: 'a poor orphan girl', 'a skull with no body' and thereafter referred to appropriately; 'she', 'the girl', 'he', 'the skull'. Thus it is clear at all times who or what is being referred to in the text. (Walshe et al. 1990:20)

Schematic structure of unsuccessful text:

Structure	**Clauses**	**Text: Aboriginal Skeletons and Skulls**
Thesis (statement)	1	All around the world the museums do need some skeletons and skulls
Argument 1 (statement)	2	Well the Aboriginals gave them some of their grandparents to put in the museum's (sic)
Argument 2 (recount)	3	Well Loir Richards is an Aboriginal
	4	and she said that some people say
		that Aboriginals have not got any feelings
Conclusion	5	The skeletons and skulls should go back
	6	where they come from and remain
Argument 3	7	You would not like it
	8	if they took your grandparents skeletons and skulls.

Commentary on text:

... the structure of Text 2 can be summarised as follows; a general statement, which could be generously interpreted as a thesis, following by a second statement, rather than an argument. Next is a short recount rather than a second argument. ... followed by the writer's conclusion regarding this topic and then there is a follow up argument. It is the conclusion that gives the clearest indication that the writer intended the text to be an Exposition. Other stages in the text are not those of a successful Exposition. ... Development of topic in Test 2 is unsatisfactory. While the text is loosely cohesive around the topic of museums and skulls and skeletons, there are problems in that none of the arguments follow logically from one another. ... There are also minor problems with referencing. The opening statement refers to 'the museums'. Such reference is unclear as the reader is not informed which museums are being referred to. ... Thus Text 2 is assessed as unsatisfactory on all three criteria. (Walshe et al. 1990:21)

Outcomes:

1. The innovation had an overall positive response from participating teachers, teachers giving an overwhelmingly positive response to the Package;
2. The innovation had a beneficial impact on students' writing. A comparative analysis of texts from Package and non-Package schools indicated that, in terms of the evaluative criteria identified by the researchers, students in Package schools produced a greater range of factual texts, and produced them more successfully;
3. Teachers' classroom practices changed as a result of participation in the innovation. While teachers incorporated ideas from the package into their teaching, they adapted these significantly to suit their needs;
4. The objectives of the innovation were largely met;
5. The model of inservice upon which the innovation is based has a number of distinctive features which contributed to its effectiveness. These included the balance of theory and practice, the demonstration lessons, and the cyclical nature of the input. (Nunan, 1992:207-208)

Case Study 2: South Australian Needs Assessment Procedures Project (SNAP)

The South Australian Needs Assessment Procedures project was initiated to develop a systematic procedure for diagnosing the strengths and weaknesses of students in their control over the spoken and written genres which they would need in order to succeed in school. The project also took on the extremely ambitious task of attempting to develop a procedure for quantifying qualitative information. My role in this project, over a two year period, was as consultant and adviser to the team and the principal researcher, Lexie Mincham.

I have included examples of two different genres, an explanation and a recount to illustrate the way in which the functional language model underlying the project has been used to provide a set of explicit criteria for evaluating students' writing. It is not difficult to see how the pro formas could also be used for instructional purposes. One of the key principles behind learner-centred instruction is that teachers be explicit about what learners are supposed to do and why, and so, in a learner-centred classroom, these pro formas could be extremely useful in helping the students see what they need to do in order to produce acceptable texts of various kinds.

Example 1

Written Language Assessment Activity : Years 8 - 10 **Explanation**				
An explanation is a factual text used to explain the processes involved in the evaluation of natural and social phenomena, or how something works. Explanations are used to account for why things are as they are, focusing on causal relations. In the school curriculum, explanations are often found in Science and Social Studies				
Name of Student Name of School Topic of explanation	Year Level/Class Date Teacher			
Criteria (Tick appropriate box)	Very Competent	Competent	Limited Competence	Not Competent Yet
Ability to carry out the task *Did the student* - write an explanation with minimal support ***Structure and organisation:*** - introduce the issue/make a general statement - use logical, sequenced explanation of how/why something occurs - use paragraphs ***Language features*** - use appropriate subject vocabulary - use appropriate tenses - use passive voice (optional) - use relevant linking words, e.g. if, when, because, consequently, since - express relationships between concepts, particularly time and cause/effect relationships - use nominalised processes - develop themes logically and consistent with explanation - use language to maintain appropriate tenor - support text with diagram (optional) ***Accuracy*** - use grammar accurately e.g. word order, verb endings, pronouns - spell and use punctuation accurately				
General comments				
Global rating (circle)	Lowest 1__2__3__4__5 Highest			

Example 2

Written Language Assessment Activity : Years 8 - 10 **Recount**				
A recount relates a series of events The focus of a recount is on events rather than on character development and plot as in narrative Recounts can be personal (retelling events in which the writer has been personally involved e g an excursion); factual (recording details of an incident/event e g news report or historical account), or imaginative (retelling events from an imaginary point of view (e g "A day in the life of a Roman slave" or "How radium was discovered)				
Name of Student Name of School Topic of explanation	Year Level/Class Date Teacher			
Criteria (Tick appropriate box)	Very Competent	Competent	Limited Competence	Not Competent Yet
Ability to carry out the task *Did the student* - write a recount with minimal support *Structure and organisation*: - provide an orientation, establishing *who* was involved, *where* and *when* the events happened - provide a sequence of events in chronological order - use paragraphs - provide a re-orientation and/or personal comment (optional) *Language features* - focus on individual participants e.g the San Francisco earthquake, Marie Curie - focus on past tense e.g simple past - use a range of action verbs e.g erupted, discovered, worked - use a range of temporal and other connectives, e.g first, then, finally, because, however, although, as well - use pronoun reference, e g it, she, this - use specific vocabulary appropriate to the information being recounted *Accuracy:* - use grammar accurately e.g word order, verb endings, pronouns (*she, her, hers, they, their* etc. - spell and use punctuation accurately				
General comments				
Global rating (circle)	Lowest 1__2__3__4__5 Highest			

BEST COPY AVAILABLE

Discussion

Thus far, I have described some of the advantages of adopting a functional model of language for assessing students' writing. I have illustrated the approach with reference to a large scale evaluation project, and also with reference to the development of some innovative diagnostic instruments for assessing students writing in key areas of the curriculum. However, I would not like to convey the impression that the approach is unproblematic, and in this section, I would like to deal with two particular problem areas.

The first of these relates to how one might compare assessment procedures based on a functional view of language with more conventional procedures. This is the problem which my colleagues and I encountered in our evaluation of the disadvantaged writing project. On one hand we wanted to adopt evaluative procedures which were fair to the innovation being evaluated. On the other hand, we did not want the evaluation process to lay itself open to charges of bias in favour of the functional approach to the teaching of writing. This dilemma relates to what Beretta calls 'program fair evaluation'. In selecting assessment procedures and instruments, the evaluators need to ensure that one of the assessment procedures or programs being investigated is not discriminated against. In his paper, Beretta provides the following examples which fail the test of being program-fair:

> Asher (1972) and Asher, Kusudo and de la Torre (1974) investigated the effect of the Total Physical Response (TPR) method compared with a "regular" program. In the 1972 report, one of the stories used in classroom training in the TPR group is presented as an example; it is entitled "Mr Schmidt goes to the office." Later in the report we are informed that one of the criterion measures used to compare experimental (TPR) and control (regular) groups is a listening test involving a "story entitled 'Mr Schmidt goes to the office'" (p.136). In view of this, it is hardly astonishing that the experimental students dramatically outperformed controls (p = .0005). (Beretta 1986:432)

In the case of the Disadvantaged Schools Project, we made the decision to adopt assessment criteria derived from the functional grammar, even though this left us open to the charge of bias. We did so, because one of our terms of reference was to determine the extent to which the innovation had an impact on students' writing. No doubt, had we embraced alternative criteria (such as evidence of 'creativity') the outcome may well have been different.

The second problem related to the so-called product-oriented bias of genre-based pedagogy. Certain proponents of "process" approaches to the development of writing have argued that genre-based pedagogy takes a normative approach to the production of texts, and focuses on the end product, the destination, as it were, rather than the route. (Some proponents of genre-based teaching, of course, have argued that process approaches focus on the route, and ignore the destination).

> Genre theory grounds writing in particular social contexts, and stresses the convention-bound nature of much discourse. Writing, therefore, involves conformity to certain established patterns, and the teacher's role is to induce learners into particular discourse communities and their respective text types. By contrast, the process approach extols individual creativity, individual growth, and self-realisation, and the teacher's role is that of 'facilitator' rather than 'director'. (Bamforth 1993:94)

Bamforth goes on to point out that the process versus product debate represents a false dichotomy, and that certain individuals on either side of the debate have taken up positions that are ideological rather than empirical. He points out, quite rightly in my view, that ultimately "the central issues of freedom and control are not alternatives between which a choice has to be made. They are really interdependent, and effective writing pedagogy will call upon both approaches."

The process product debate has also suffered from confusion between syllabus design and methodology. To my mind, the strength of the genre approach rests on the principles it sets out for the selection of content. This is essentially a syllabus design issue. The process approach, on the other hand, is oriented towards classroom action, and its concerns are therefore essentially methodological. Any comprehensive approach to pedagogy must incorporate syllabus design, methodology, and assessment.

In this section, I have looked at two criticisms which have been made of the functionally-based genre approach to pedagogy. The first of these is in identifying methods of evaluating genre-based curricular innovations against what, for want of a better term, we might call 'traditional' methods of teaching and assessing written language. This first criticism is in the nature of a 'straw person' argument, in that it can be applied to any approach to pedagogy, when the purpose of the assessment is to evaluate whether or not the principles underlying the innovation are actually reflected in the written (or oral) production of the students. The second criticism relates to the product oriented bias of the adoption of a genre approach. Once again, I find the argument fundamentally flawed for the following reasons. In the first place, the argument that product oriented approaches somehow stifle the creativity of the writer overlooks the fact that creativity has to be measured against something, and that something is generally taken to be a set of conventions or 'rules'. In any field in which 'creativity' is given a look-in, whether it be painting, creative writing, or the production of academic discourse, the creative artist must, in the first instance, master the conventions of the discourse. In other words, one must master the rules in order to transcend them.

Conclusion

In this paper, I have tried to show the potential of adopting a functional approach to linguistic analysis for assessing written texts. I have argued that functional grammarians have, over the last few years, provided language teachers with criteria for evaluating the extent to which learners have gained control of the grammatical and discoursal features of genres which are highly valued within academic contexts. I have tried to illustrate the potential utility of the approach by describing two innovations from either end of the curriculum continuum. The first is a large scale evaluation project within the primary school sector in Sydney, Australia. The second is a needs assessment and evaluation project carried out in South Australia. While the purpose of the paper has been to advocate the adoption of a functional approach to the assessment of second language proficiency, it has not done so uncritically. I have also discussed two criticisms which have been made of the approach.

References

Bamforth, R. (1993). Process versus genre: Anatomy of a false dichotomy. *Prospect*, 8 (1 & 2):89-99.

Beretta, A. (1986). Program-fair language teaching evaluation. *TESOL Quarterly*, 20(3):431-444.

Brown, J.D. (1989). Language program evaluation: a synthesis of existing possibilities. In R. K. Johnson (ed.) *The Second Language Curriculum*. Cambridge: Cambridge University Press.

Halliday, M.A.K. (1985). *An Introduction to Functional Grammar*. London: Edward Arnold.

Mincham. L. (1991). The SNAP Project. Department of School Education, South Australia.

Nunan, D. (1992). *Research Methods in Language Learning*. Cambridge: Cambridge University Press.

Nunan, D. (1993). *Introducing Discourse Analysis*. London: Penguin.

Walshe, J., Hammond, J., Brindley, G. and Nunan, D. (1990). *Evaluation of the Disadvantaged Schools Writing Project*. Sydney: National Centre for English Language Teaching and Research.

STEPPE BY STEP: A CULTURAL APPROACH TO LANGUAGE TASKS

Colin Barron

Introduction

At one end of the ESP spectrum we have the Spack model, which is independent of the serviced discipline:

> To initiate students into the academic discourse community, we do not have to change our orientation completely, assign tasks we ourselves cannot master, or limit our assignments to prescribed, rule-governed tasks. We can instead draw on our own knowledge and expand the knowledge and abilities of our students. (Spack 1988:47)

and:

> The materials we use should be those we can fully understand. The writing projects we assign and evaluate should be those we are capable of doing ourselves. (Spack 1988:41)

At the other end of the ESP spectrum is the Widdowson model which is fully mimetic of the serviced discipline:

> ESP is (or ought logically to be) integrally linked with areas of activity (academic, vocational, professional) which have already been defined and which represent the learners' aspirations. The learning of ESP is in consequence an essentially ***dependent*** activity, a parasitic process, and it follows that the pedagogy of ESP must be dependent too. It has no purposes of its own; it exists only to service those that have been specified elsewhere. (Widdowson 1983:108-109; emphasis in original)

Most ESP teachers will probably feel sympathetic towards one of these two positions. Both views are, I believe, untenable because they are alteric and mimetic, although in different ways, and are based on a theory of deficiency. They assume language to be an independent system, a tool, which can be described structurally and functionally and taught independently of social, cultural and political factors. The concept of social and cultural awareness is present in the task, an important feature of many ESP courses.

In this paper I am not going to take up a position in the middle of the spectrum between these two unidimensional models. ESP will be placed firmly within the multidimensional space that constitutes the students' chosen disciplinary culture, not as an adjunct that either mimics the disciplinary culture, or provides an alternative. The multidimensional space includes social, cultural and political factors as well as functional ones. Socio-cultural factors are primary variables in ESP because it is a part of education whose purpose is to produce competent members of society and because language is a part of a unified theory of social action. "Language is not acquired without culture" (Ochs 1988:38), so that socialisation through language should be an important factor in ESP. The position adopted was realised through a project designed for Computer Engineers.

ESP needs a model of the part language plays in the socialisation of novices into their disciplinary cultures; a perspective that looks at the multidisciplinary nature of disciplinary cultures, one that sees it as a unified system of social action linking events, actions and processes with linguistic and generic features. A multi-dimensional space is complex, so I will provide a map to guide the ESP teacher through this space, and to show the way forward to answering the questions posed by Spack:

1. How much responsibility do we have in assisting to initiate the students into the cultures of their chosen disciplines?
2. How much knowledge of the serviced discipline is needed by the ESP teacher?
3. What are the most appropriate sources of materials for the students?

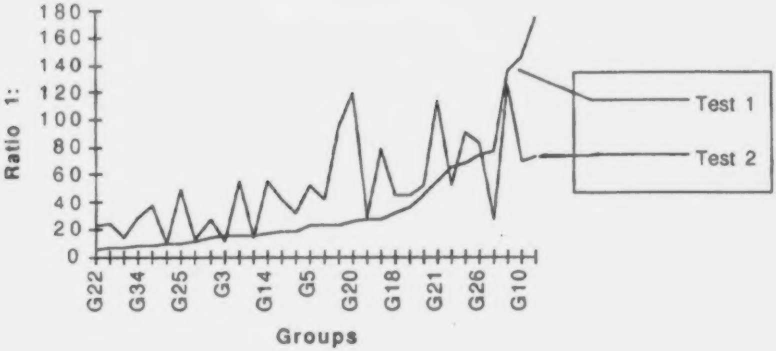
The Project

The project is called "Design and Make" and is the fourth and last in a series of projects for the first-year computer engineering and computer science students at the University of Hong Kong. The project is quite simple in design. In it the students design and make a crane out of paper and sellotape only. They are provided with a wooden base on which to place their crane. Each class is divided into groups of four students or so. Each group designs a crane according to the specifications given (see appendix). They then make the crane. No limit is given on the amount of paper and sellotape that each group may use, but weight is an important consideration because the completed cranes are rated according to their mass: applied load ratio.

The cranes are tested in one of the engineering laboratories by applying weights to the ends of the cranes until they collapse or deform markedly from the vertical. The tests are conducted twice so that the students can learn from the lessons of the first test and redesign their models to improve their results. An engineering technician is available during both tests to explain to the students how and why their cranes fail. They then use this information in their written reports. The results of all the groups are collated and given to the students so that they can compare their results with the rest of the class, and so they can determine the optimum value. The ratios ranged from a low of 1:5.7 to a high of 1:174, i.e. the latter crane bore 174 times its own weight before collapsing. Most of the groups improved their result on the second test, some considerably so. The few groups which did not, achieved a result only marginally worse (fig. 1).

Figure 1

Project 4 Results



The handout for the project consists of one sheet and this generates four to five weeks of work. Supplementary handouts on report-writing are given during the project, but even these are kept to a minimum because the students are expected to determine what goes in their reports based on what they do, how they do it, why they do it, and their results.

The entire course was evaluated at the end of the year. The design and make project proved to be very popular with the students, even though it was not directly related to their discipline. Table 1 shows the relative popularity of the projects in 1993.

Table 1

Students' preferences for projects (N = 99)

Project	**1**	**2**	**3**	**4**	**Press dossier**
1 (best)	11	7	37	41	0
2	18	21	25	21	2
3	26	13	13	15	9
4	19	21	13	8	20
5 (worst)	11	13	5	3	25

The table shows the actual number of students who rated the projects from best to worst. Projects 3 and 4 were clearly the most popular, with 42.7% of those who responded putting the design and make project (project 4) first, i.e. as the best/most interesting. A further 24.1% of respondents placed it second, i.e. nearly 70% of the students rated the project very highly. Project 3 required the students to choose their own topic. Projects are rated very positively if their objectives are syntonic with the expectations of the students' chosen disciplines and also require students to make a large contribution.

The comments made by students on their reasons for liking or not liking the design and make project are instructive. This project attracted more positive comments than the other projects. It also attracted fewer negative comments than the others. Question 26 on the evaluation form was an open-ended one which asked "Which were the best/most interesting parts of the course?" The design and make project 4 produced the largest number of responses to this question. Question 27 asked "How could this course be improved?" Several replies were "more projects like no. 4."

Both the positive and negative comments on the project are revealing. The positive comments stress relevance, cognitive development, co-operation and socialisation (fun), e.g.:

> relevance - "appropriate for engineering"; "useful when technical reports needed"
>
> the fun aspect - "like a game"; "learning many things but doing little"; "light work load". This project was deliberately placed at the end of the course for these reasons, because at this time of the academic year the students are overloaded with work in their other subjects. So the project satisfied that criterion - language as social action.
>
> cognitive development - "stimulates thinking"; "learn something other than English"; "learning many things but doing little". Again these were objectives of the project. Language cannot be acquired without culture.
>
> co-operation - "teamwork"; "independent work". I am not sure what this means, but I take it to mean independent of the teacher since all the students worked in groups. Another objective of the project. ESP/EAP is not parasitic, mimetic or alteric.

The negative comments were far fewer and concentrated on irrelevance and technical matters:

> relevance - "irrelevant to CE" (computer engineering)
>
> fun - "wastes time"

> cognitive development - "can use existing data". I am not sure what this means as this was the first time that the project had been done, so no data existed.
>
> co-operation - "material written by non-technical personnel". This is interesting because the handout was written by an engineer.

The comments indicate the awareness that computer engineering and computer science students at the University of Hong Kong have of the relevance of tasks and the relationship between language and their disciplinary culture and their ethnic culture.

Task

I am using the notion of task here as a unit of social action. We need to distinguish between "task" and "activity". Both are used indiscriminately in the literature to refer to the kind of project we are discussing here. Long (1985), Nunan (1988) and Mohan and Smith (1992) all use the term task. Mohan and Smith define it as a common unit of "both language teaching and learning, and content teaching and learning" (Mohan & Smith 1992:82). Task is also an important unit of activity in work, and an important unit in culture and socialisation in anthropology, e.g. Lave's (1977) study of tailoring in Liberia, and Childs and Greenfield's (1980) study of Zinacanteco weavers in Mexico, where adult task-based talk varied with the specific weaving task and the level of skill of the child (novice).

Mohan (1986) and Ochs (1988) use the term activity in the same sense. For Ochs, "activity is both a behavioral unit, in the sense of a sequence of actions associated with particular motivations and goals, and a process, in the sense of praxis" (1988:14). For Mohan, "an activity means a social practice, a form of social life that has a publicly acknowledged structure and standards" (1986:48), "a combination of action and theoretical understanding" (1986:42).

Language teachers have tended to adopt the behavioural, functional definition of a task as a unit whose aim is its successful completion. For example, Long defines a task as:

> a piece of work undertaken for oneself or for others, freely, or for some reward ... examples of tasks include painting a fence, dressing a child, filling out a form, buying a pair of shoes, making an airline reservation, borrowing a library book, taking a driving test, typing a letter, weighing a patient ... In other words, by "task" is meant the hundred and one things people do in everyday life, at work, at play, and in between. (1985:89; quoted in Nunan 1988:85)

Nunan (1988:85) criticises this functional notion of the task in the task-based syllabus developed by Long and Crookes (1986) because it is unitary:

> the specific nature of the task and the content on which it is based are unimportant ... as long as learners are productively engaged in a task, (so that) they will be acquiring the target language.

A task is not an end in itself. It is a unit of social action in the process of constructing knowledge as an individual becomes a member of his/her culture. A task is socially determined but individually interpreted, because "for any social behavior, there may be different *realms* of understanding that experience" (Ochs 1988:7; emphasis in original). It is a means by which knowledge is exchanged and a novice is socialised into becoming a member of the culture. Language and socio-cultural knowledge are interdependent (Ochs 1988:14). The specific nature of the task and the content are important because "language is not acquired without culture" (Ochs 1988:38). We also need to avoid the ethnocentric pitfall of determining and judging tasks from the viewpoint of a particular culture, be it ethnic or disciplinary. Tasks we give the students should therefore be those that are important in their cultures.

I therefore wish to propose differences between activity and task.

1. The aims of an activity are short-term, its successful completion, and it may or may not be part of a task, while a task has long-term aims - the maintenance of the culture, and consists of one or more sub-tasks.

2. A task is socially constructed, while an activity is often individually constructed.

3. Tasks have the purpose of initiating novices into the culture through their participation in joint tasks with more knowledgeable persons. Tasks are thus critical for the development of cognitive skills (Ochs 1988:15).

4. Task is an important concept in education because the purpose of education is to produce competent members of society. An educational task may consist of one or more skills or activities, but its purpose is not the completion of the skills or activities per se, but the maintenance of the culture. It is thus part of a much bigger system, a system of delayed exchange whereby the experts pass on their knowledge to initiates, who in turn pass on their knowledge later, but not necessarily in exactly the same form. The return is social rather than functional, the satisfaction of knowing that the culture will be maintained through an active and dynamic process of socialisation.

Discussion

The crane project is a task that is socially determined by the teacher but individually interpreted by the students. It is determined by the teacher in order to give the students practice in problem-solving in engineering, specifically in invention, one of the three kinds of engineering problems (Higgins et al. 1989:171). This kind of problem determines the method of solving it, the STEPPE method (Higgins et al. 1989:172), which consists of:

1. State the problem
2. Think about possible solutions
3. Evaluate solutions and choose one of them
4. Present the chosen solution
5. Produce the chosen solution
6. Evaluate the process of problem-solving

But the project is individually interpreted in that the students are free to interpret the problem in any way they wish. Individual interpretation is shown most clearly in the variety of designs and models of crane that they produce. But in addition, the students interpreted the instructions rather more freely than we had intended. Some groups interpreted the instruction to use only paper and sellotape to their advantage by using sellotape as a structural member. Since sellotape has a much greater tensile strength than paper, their results on the first test were very impressive. This can be seen on figure 1 above where the results of the first test suddenly take off at about 1:30.

The objective of the project is epistemological and social, a unit of action in the process of the students' construction of knowledge in engineering. But it is not mimetic. It deliberately does not imitate anything in the students' engineering course, although it is similar to other tasks they do. Its purpose is to take the students through analytical and socialisation processes to arrive at a synthesis which exemplifies one type of problem-solving in engineering, invention. Research into disciplinary cultures shows that different disciplines have different key concepts. For example, "in fundamental-particle physics, solid-state physics, and molecular biology, discovery is a central notion... whereas in mechanical engineering the concept is largely replaced by that of invention" (Becher 1984:178-179).

Although the project is socially determined, it is very successful in generating a considerable amount of discussion and comment. So the students determine the outcome of the task and regulate the considerable amount of social interaction that takes place. This is evident from their comments in the evaluation of the project; "like a game", "teamwork", "stimulates thinking".

The project exploits the task as a unit of socialisation for the students into their disciplinary culture because neither the task nor the product, the written report, are ends in themselves. Socialisation is the third element, a necessary element, which links the analytical level of determining the linguistic requirements of a report with the synthetic level of satisfying the objective of the task, practice in the problem-solving technique of invention. This element is the process through which knowledge is constructed. Socialisation determines the outcomes and thus establishes power relationships. Socialisation is quite simply, in Ochs' definition, "the process by which one becomes a competent member of society" (1988:5). In the crane project the students socialise through language ("like a game"), and socialise to use language ("teamwork"). It is one step on the way to enabling the students to become competent members of the engineering culture, in this case through one of the key concepts in engineering, invention.

The use of terms such as "step" and "level" may indicate that the socialisation model operates at three discrete levels. This is not the case. The three steps or levels, that of the task or the cultural level, the socialisation level and the analytical level, are mutually interdependent and operate at the same time. The three have been isolated merely to make them explicit. The model is, like a map, a two-dimensional representation of a multidimensional entity which is much more complex.

The crane project accepts that the students (novices) have an important part to play in the culture and recognises this by encouraging socialisation between all members, the students, the English teachers and the engineering staff. In this project the latter are the laboratory technicians. Its aim is not alteric. It does not set up a goal, invention, to be reached with little or no active interaction between the students and their teachers, i.e. which bypasses the socialisation process by establishing a direct line between the analytical level and the goal. Nor does it set up an opposition between engineering and language. Instead, the project recognises, as anthropologists have done for some time, that all participants need to be active contributors, including the novices (the students). "The novice is not a passive recipient of socio-cultural knowledge but rather an active contributor to the meaning and outcome of interactions with other members of a social group" (Schieffelin & Ochs 1986:165). The students recognise that they are active participants in the engineering culture and that the crane project is making a positive contribution to their acquisition of the culture of engineering in a multidimensional way, as the students recognised: "stimulates thinking", "learn something other than English," "teamwork". The students are recognised as being members of the culture, although novices, hence their positive comments.

The experts (the English and engineering teachers) are needed to determine the task and to be active contributors to the meaning and outcome of interactions. In Vygotskyian terms, the socially determined task is necessary because it:

> emphasizes the role of more knowledgeable members in facilitating learning. Novices are able to carry out particular tasks through "guided interaction"; they develop skills in a "zone of proximal development" as

> they move from guided or collaborative to independent action. Within this framework, cultural knowledge both organizes and is acquired through these communicative activities. (Schieffelin & Ochs 1986:166)

The crane project establishes a multidimensional space in which all those involved have a place. Mimesis can be avoided by the experts being allocated different tasks. They can complement one another, and the language tasks can be those that the engineering teacher does not have time to do. This avoids mimesis, i.e. imitating exactly what is taught in the students' classes. It also avoids alterity, which places the language teacher outside the students' target culture, creating separate power bases which may fracture or ignore the relationships between the two kinds of "experts" and their novices, and encouraging language to be considered as a separate system.

The traditional disjunction between language teachers and subject teachers cannot continue because language "is continuous with other dimensions of social action, rather than constructing a privileged domain of commentary on them" (Merlan & Rumsey 1991:223). It cannot continue because knowledge is an abstract entity that is given dynamic form by language. The construction of knowledge is dependent on the social dynamic between individuals because a single individual cannot carry the whole body of knowledge of a culture, it is just too diverse and complex:

> Bodies of knowledge, structures of understanding, conceptions of the world, and collective representations are extrinsic to any individual and contain more information than any individual could know or learn. Culture encompasses variations in knowledge between individuals but such variation, although crucial to what an individual may know and to the social dynamic between individuals, does not have its locus within the individual. (Ochs & Schieffelin 1984:284)

The social dynamic is achieved through socialisation between members of the culture. The ESP teacher provides the linguistic "means of access to particular domains, or spheres of social action" (Halliday 1993b:45). The subject teacher provides the conceptual means of access.

Since culture does not reside within any one individual, but is within the collective body of members of the society, the need for members of the society to share their knowledge and experiences through socialisation with each other is essential. This follows the Vygotskyian principle that "higher mental functions are social before they are internalized by the individual, and that they become internalized by means of social interactions" (Pelissier 1991:81). Everybody in the society needs to take part in the social interactions and thus be considered a member of the culture, whether expert or novice. The socialisation model suggested here denies the traditional disjunction between subject and language teachers, with the students caught in the middle:

> We cannot work with the traditional disjunction ... in literacy theory, according to which there are two human populations, the writers and the non-writers, with a clear dichotomy between them. Rather we should envisage a multidimensional space within which human beings construe experience, with certain regions within that space being taken up preferentially by different groups to resonate with the different conditions of their material existence. (Halliday 1993a:13)

The ESP teacher has to be located within the multidimensional space that is the students' target culture, facilitating the students to enter the culture by providing tasks that resonate with it. This entails co-operation between all who are involved - content teachers, language teachers and students. In the crane project this co-operation is exemplified by the students carrying out their tests in one of the engineering laboratories with a technician and the language teacher both present. The technician's role is to explain the technical reasons for the collapse of the cranes, which the students incorporate into their written reports. The language teacher learns along with the students and is in a position to assess the content of this part of the report.

This is very different from the Spack model in which the concepts of the humanities culture are reproduced in the students and which places the ESP teacher outside the multidimensional space that the students are entering. "Students can learn general inquiry strategies, rhetorical principles, and tasks that can transfer to other course work" (1988:40-41). There is no multidimensional space envisaged here, just a single essentialist level determined by the teacher, based on his/her humanities background. It is here that the problem appears in Spack's argument because, while rhetorical principles may be transferable because they are at the analytical level, tasks do not transfer to other course work because they are at the cultural level; they are socially constructed. Spack admits that, "each discipline offers a different system for examining experience, a different angle from looking at subject matter, a different kind of thinking" (1988:38, citing Maimon et al. 1981). Becher has analysed the differences between academic disciplines. "In summary, the three disciplines [history, sociology and physics] display fundamental differences not only between types of evidence and procedures for proof, but also in the ways in which others' work is evaluated and in the modes in which arguments are generated, developed, expressed and reported. All these have implications for the underlying knowledge structure" (1987:273). Bazerman (1988:82-84) explains how physics has deliberately fostered a distinct cultural and rhetorical identity since the time of Newton. Humanities tasks do not resonate with engineering tasks, for example, so they do not help in initiating engineering students into their chosen disciplinary culture. So the Spack model is alteric because it sets up an opposition between the ESP teacher's disciplinary culture and that of the students chosen discipline.

The Spack model is mimetic because two cultures are imitated, but they are cultures socially determined by the ESP teacher - the teacher's disciplinary culture, the humanities, and his/her ethnic culture, specifically "background knowledge that readers are expected to bring to written texts (e.g. knowledge of American history)" (Spack 1988:43). We should bear in mind Foucault's admonition:

> what entitles its (the West's) culture, its science, its social organization, and finally its rationality itself, to be able to claim universal validity: was this not a mirage associated with economic domination and political hegemony?" (quoted in Young 1992:9)

One result of the Spack model is cultural imperialism and the continued hegemony of the ESP teacher over the students because the students are neither novices of the teacher's ethnic culture nor his/her disciplinary culture.

The Widdowson model is clearly mimetic because it encourages parasitism. The ESP teacher is grafted onto the disciplinary culture, and is not a part of it. Language is considered to be an independent system (Bazerman 1988:155). It is a tool, it is neutral and thus subservient to the disciplinary culture. As a result, language is marginalised.

This model has no place for the ESP teacher in the initiation of the students into their disciplinary culture because the ESP teacher is squarely outside the culture. At this end of the ESP spectrum, the Widdowson model creates an independent, neutral acultural world with little relevance to the students' worlds or that outside the classroom. The result, as in the Spack model, is cultural imperialism because the culture is that of western science and engineering:

> I assume that the concepts and procedures of scientific inquiry constitute a secondary cultural system which is independent of primary cultural systems associated with different societies. So although for example, a Japanese, and a Frenchman, have very different ways of life, beliefs, preoccupations, preconceptions, and so on deriving from the primary cultures of the societies they are members of, as scientists they have a common culture. In the same way, I take it that the discourse conventions which are used to communicate this common culture are independent of the particular linguistic means which are used to realize them. (Widdowson 1979:51)

Rose and Rose have demonstrated that the neutrality of science is a myth most obviously because "a paradigm is never value-free. A paradigm is never neutral" (1980 [1971]:28). Barron (1991:174-176) has argued that science and engineering are part of the primary culture. The Widdowson model is thus alteric, the opposition being between the subject teacher within the disciplinary culture and the ESP teacher outside it, and between Western culture and the cultures of the students. There is a direct route from the analytical level to the disciplinary cultural level, completely bypassing the socialisation level. The students are caught between the two levels. Since they have little or no chance to socialise in their discipline through language, they naturally opt for their chosen discipline, so marginalising language. They may also be caught between their culture and the western culture and they have to resolve this disjuncture, not always successfully. The model is a unidimensional space, rather than a multidisciplinary, multidimensional one.

The mimetic and alteric qualities of the Spack and Widdowson models, aping the West and setting up oppositions between the ESP teacher and the chosen disciplinary cultures of the students are educationally, culturally and politically unacceptable. They are educationally unacceptable because ESP is a part of education whose aim is to produce fit members of society. They are culturally unacceptable because they do not provide the students with tasks that are resonant with their disciplinary culture, and because they deny the opportunity to interact at the socialisation level. They create a route leading directly from the analytical level to the cultural level, resulting in cultural imperialism in which cultures are presumed to be similar. Unfortunately Halliday's wish is not yet a reality: "I hope by now we are beyond the point where we have to pretend that everybody's world view is alike" (1993a:11). Spack's and Widdowson's models recognise a single space, in which cultures do not work together and learn from each other. Cultures are dynamic systems that individually interpret input from others as they change. We need to create multidimensional systems in which different cultures work together.

The models are politically unacceptable because mimesis and alterity reinforce the hegemony of a dominant primary culture, which includes its science and engineering, and do not fulfil the educational objectives of producing fit members of society, i.e. those who are independent, and of maintaining the culture. Instead, they socially construct "an ideological common sense which holds for everyone" (Fairclough 1989:86). Individual interpretation does not take place because "the materials we use should be those we fully understand" (Spack 1988:41). There is no room in this model for any interpretation by either the teacher or the students. Materials should be "drawn from a specific field, if the area of study is one that the instructor is well versed in" because they allow instructors to avoid placing themselves in the awkward position of presenting materials they do not fully understand" (Spack 1988:43). If the teachers fully understand the materials, the amount of socialisation with the students will be considerably reduced because the channel of communication is one-way. This is a static world constructed entirely by the teacher. A world in which materials are objects, not discourses. A world in which cultures are static, immutable objects, not dynamic entities that are constantly changing.

The materials we use should be dynamic discourses, not static objects, constructed socially within the target culture, but allowing for interpretation by each individual student. They should be those in which participants are not dichotomised, language is not abstracted out for emphasis, and events and actions are not marginalised. Events, actions and language are all part of a multidimensional exchange system in which everybody participates. Teachers, both language teachers and subject teachers, and students should be able to explore the multidimensional space of the two cultures together, each providing their own level of expertise of events, actions and language in a unified approach to social action.

Conclusion

The crane project exemplifies the multidimensional nature of disciplinary cultures, one in which all participants have important roles to play, including ESP teachers. This paper has provided a map to guide ESP teachers through this multidimensional space so that they can play their part in initiating the students into their disciplinary cultures. Like any map, it is a two-dimensional representation of a multidimensional entity and therefore cannot show the true complexity of the reality. The map shows three steps, or levels: the analytical level, which includes the linguistic and generic realisations of the cultures, the socialisation level where knowledge is constructed between all members of the cultures, and the cultural level, which consists of tasks that are socially constructed and which exemplify key concepts of the cultures.

ESP's role in the initiation of students into their disciplinary cultures is complex, consisting of several facets, educational, linguistic, cultural and political. The model outlined here does not assume simplistic oppositions between the ESP teachers' disciplines and those of the students, nor a "common-sense" view of the world. Neither does it copy exactly the practices of the students' chosen disciplines. It encourages the construction of knowledge through socialisation, leading to the realisation of key concepts that are typical of the disciplinary culture, via the analysis of linguistic and generic features. Suppression or denial of socialisation leads to repression of the imagination and innovation of students by prescribing materials which are alteric or which mimic those in their chosen disciplines. "Where not repressed, the mimetic faculty may serve as a tool of repression in the 'civilizing' project of Enlightenment" (Taussig 1993:254). Mimesis in education represses because it does not fulfil the educational aim of producing competent members of society. Society needs independent social actors, not mimics. Alterity represses because it leads to cultural imperialism, either dominance of the teachers' ethnic cultures, or their disciplinary cultures, or both. Mimesis and alterity both result in the marginalisation of language, and hence a confusion of separate unidimensional spaces instead of a unitary multidimensional space.

References

Barron, C. (1991). Material thoughts: ESP and culture. *English for Specific Purposes*. 10(3):173-187.

Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental article in science.* Madison: University of Wisconsin Press.

Becher, T. (1984). The cultural view. In B.R. Clark (ed.), *Perspectives in Higher Education*. (pp.165-198). Berkeley: University of California Press.

Becher, T. (1987). Disciplinary discourse. *Studies in Higher Education*. 12(3):261-274.

Childs, C. P. and Greenfield, P.M. (1980). Informal modes of learning and teaching: The case of Zinacanteco weaving. In N. Warren (ed.), *Studies in cross-cultural psychology.* Volume 2 (pp.269-316). London: Academic Press.

Fairclough, N. (1989). *Language and power.* London: Longman.

Halliday, M.A.K. 1993a. New ways of meaning: A challenge to applied linguistics. In M.A.K. Halliday. *Language in a changing world* (pp.1-41). Occasional paper no. 13. Applied Linguistics Association of Australia.

Halliday, M.A.K. 1993b. The act of meaning. In M.A.K. Halliday. *Language in a changing world* (pp. 42-61). Occasional paper no. 13. Applied Linguistics Association of Australia.

Higgins, J.S., Maitland, G.C., Perkins, J.D., Richardson S.M. and Warren Piper, D. (1989). Identifying and solving problems in engineering design. *Studies in Higher Education* 14(2):169-181.

Lave, J. (1977). Cognitive consequences of traditional apprenticeship training in West Africa. *Anthropology and Education Quarterly.* 8(3):177-180.

Long, M.H. (1985). A role for instruction in second language acquisition. In K. Hyltenstam & M. Pienemann (eds.), *Modelling and assessing second language acquisition.* Clevedon: Multilingual Matters.

Long, M.H. and Crookes, G. (1986). Intervention points in second language classroom processes. *Working Papers 5,2.* Department of English as a Second Language, University of Hawaii.

Merlan, F. and Rumsey, A. (1991). *Ku Waru: Language and segmentary politics in the western Nebilyer Valley, Papua New Guinea.* Cambridge: Cambridge University Press.

Mohan, B. (1986). *Language and content.* Reading, MA: Addison-Wesley.

Mohan, B. and Smith, S.M. (1992). Context and cooperation in academic tasks. In D. Nunan (ed.), *Collaborative language learning and teaching* (pp. 81-99). Cambridge: Cambridge University Press.

Nunan, D. (1988). *The learner-centred curriculum: A study in second language teaching.* Cambridge: Cambridge University Press.

Ochs, E. (1988). *Culture and language development: Language acquisition and language socialization in a Samoan village.* Cambridge: Cambridge University Press.

Ochs, E. and Schieffelin, B. (1984). Language acquisition and socialization: Three developmental stories and their implications. In R. Shweder & R. Levine (eds.), *Culture theory: Essays on mind, self and emotion* (pp. 276-320). New York: Cambridge University Press.

Pelissier, C. (1991). The anthropology of teaching and learning. *Annual Review of Anthropology*, 20:75-95.

Rose, S. and Rose, R. (1980). The myth of the neutrality of science. In R. Arditti, P. Brennan and S. Cavrak (eds.), *Science and liberation* (pp. 17-32). Montreal: Black Rose Books (originally published in 1971).

Schieffelin, B. and Ochs, E. (1986). Language socialization. *Annual Review of Anthropology*, 15:163-191.

Spack, R. (1988). Initiating ESL students into the academic discourse community: How far should we go? *TESOL Quarterly*, 22(1):29-51.

Taussig, M. (1993). *Mimesis and alterity: A particular history of the senses.* New York: Routledge.

Widdowson, H.G. (1979). *Explorations in applied linguistics.* Oxford: Oxford University Press.

Widdowson, H.G. (1983). *Learning, purpose and language use.* Oxford: Oxford University Press.

Young, R. (1992). *White mythologies: Writing history and the West.* London: Routledge.

COHERENCE AND CONTINUITY IN THE TASK-CENTERED LANGUAGE CURRICULUM: GLOBAL EDUCATION AS A FRAMEWORK FOR TASK-BASED LANGUAGE TEACHING

Brenda Bushell and Brenda Dyer

Introduction

Although the idea of using the learning "task" as a basic planning tool for second/foreign language teaching is not a new one, there still remains confusion about the place of tasks within a curriculum. Nunan (1989a) notes that learning tasks, seen as an integrated set of processes within the curriculum, can provide an appropriate language learning basis for a variety of learner types from children to adults within a range of situations from general purpose English programs to English for Academic purposes (EAP). In this paper, an examination of the learning needs of a population of 36 students (2 freshmen classes) studying English for Academic Purposes (EAP) at International Christian University (ICU) in Japan is undertaken, and a task-based language syllabus in global education is outlined.

Established shortly after World War II, ICU began as a language institute, then expanded by developing a college of liberal arts programs to provide international education within a bilingual environment. Today ICU has an enrollment of approximately 2,000 students, 500 of whom are freshmen taking a variety of English language courses in the English Language Program (ELP) at ICU. These courses are aimed at preparing students for upper level classes, and as preparation for writing a senior thesis in English if they so choose. To investigate the present English language curriculum of ICU in the light of students' perceived needs for studying English and the wider goals of the undergraduate program as a whole, a learner needs analysis was administered. The research question was fourfold:

1. What kind of English language study do students perceive as being important for their success and for what purposes?
2. How do students' perceived needs match up with the professors' assessment of required academic tasks, as measured by a task analysis completed by five ICU instructors in 1991?
3. What support is there for a global education curriculum within the ELP?
4. Is the Environment Unit presently in place in the ELP curriculum compatible with the results of the needs analysis?

Before examining the results of the needs survey, an overview of task-based language teaching and global education will be presented.

Long (1985:89) defends task-based language teaching on psycholinguistic grounds. According to Long, language acquisition does not occur in a linear, additive fashion, in the discrete units of language traditionally presented - vocabulary, structure, function - but rather, in "chunks of meaning." The meaningful unit of analysis and input is task, which he defines as "a piece of work undertaken for oneself or for others, freely, or for some reward." In fact, the term "task" has been defined in a variety of ways. Candlin (1987:10) defines task as a sequential and problem solving social activity, which involves application of existing knowledge to attainment of goals. Finally, Nunan (1989:10) defines a task as "a piece of classroom work which involves learners in comprehending, manipulating, producing, or interacting in the target language while their attention is principally focused on meaning rather than form."

Task-based learning provides a purpose for the use and learning of a language other than just learning language items for their own sake. The syllabus then is not a list of linguistic items, nor a description of what a learner will be able to do at the end of a course, but a list of tasks and activities that learners will do. However, a problem with focusing purely on the learning process, not the learners' individual needs or outcomes, is that the syllabus lacks logic and coherence, as was the case in Prabhu's "Bangalore Project." Although his early experiment in task-based teaching was ground-breaking, the lists of classroom tasks - information gap, opinion gap - were a random collection of unconnected activities.

Candlin (1987:20) suggests that two important factors which facilitate the comprehension of even difficult language tasks are content continuity (the extent to which the content relates to the real world interests or needs of the learner), and process continuity (the coherence, continuity and inter-relatedness of tasks). We believe the answer to achieving coherence and continuity in a task-based syllabus is to use a subject area as the basis for tasks. The content area itself, with its academic forms and demands, will provide a non-linguistic rationale for selecting and grading tasks. The content-based syllabus proposed by Mohan (1986) overlaps with Prabhu's project in its objectives: to develop language through classroom activities designed to promote cognitive skills. Of course, many content areas can provide a meaningful context for task-based language learning: literature, history, science, etc. The choice of both course designers and students was that of global education.

Just as the concept of task varies from one teaching environment to another, so has the definition of global education. In the post W.W.II years, global education emerged in America to serve the national interests of enhancing political, economic and sociocultural influence throughout the world. The purposes and goals of global education were principally to prepare students with a realist-geopolitical world view (Lamy 1983). Students were taught that military and economic power was essential in dealing with world conflict and that their nation must take the responsibility of organizing and supporting the "free world" in its conflict with the Soviet system. International education programs were also aimed at the citizens of foreign, particularly developing, nations. The goal was to ensure that these countries would develop "like us" and would remain on "our" side in the constant struggle with the Soviets (Smart 1971).

But as the Cold War ended the operational principles of competition and domination, which characterized the last century, became no longer feasible in the global village of the contemporary world.

The idea of perspective consciousness, the awareness of alternative world-view is now central in the new definition of global education as "an approach to education in which students learn to perceive and understand the world as a global system, and at the same time to recognize the benefits, costs, rights and responsibilities inherent in their participation in the system" (Lyons 1992). Education should "prepare an individual for entry into a fluid regenerating society rather than perpetuate existing conventions, roles and structures" (Goble and Porter 1977:24).

Developing a Task Syllabus: Identifying Tasks by Needs Analysis

Once one has decided on task as the basis for a syllabus, how does one identify, select and sequence tasks? Long (1985) uses a needs analysis as the point of departure. A learner needs analysis identifies "target tasks" which are those tasks required for an individual to function adequately in a particular target domain - academic or occupational. Examples of target tasks are buying a train ticket, bargaining at a market, taking lecture notes. The following is Long's procedure in moving from needs analysis to a task syllabus:

1. Conduct a needs analysis to obtain an inventory of target tasks.
2. Classify the target tasks into task types.
3. From task types, derive pedagogical tasks.
4. Select and sequence the pedagogical tasks to form a task syllabus (p. 91).

Target tasks are usually considered the "real-world" tasks; in the case of English for Academic Purposes, the target tasks will be academic tasks. Nunan (1988) differentiates between two types of needs analysis used by language syllabus designers - a task analysis and a learner analysis. An example of a task analysis can be seen in Appendix A. This survey was distributed to faculty members of ICU in 1991 to determine what academic tasks were required of Japanese students after completing the English Language Program.

Such a task analysis identifies the "real-world" (in this case, academic) language tasks which learners are preparing to undertake. The syllabus designer then decides the subordinate skills and knowledge required by the learner to carry out these tasks. Appendix B shows our questionnaire which was based on Nunan's model (1988:63) used to collect information from learners as a basis for course design. This learner analysis can guide the selection of content and methodology. It can also show areas of possible conflict between learner and teacher goals - e.g. if learners feel memorizing vocabulary lists is the most useful way to spend class time, the teacher may need to defend/negotiate a different language teaching approach. The learner analysis attempts to involve learners and teachers in exchanging information so that learning and teaching agendas are not at cross-purposes.

Summary of Task Analyses

There are several 'task-analyses' which have been done in recent years to determine what academic tasks are actually required in undergraduate courses. Horowitz (1986d) found that academic writing tasks assigned by professors fell into seven categories: summary of/reaction to a reading, annotated bibliography, report on an experience, connection of theory and data, case study, synthesis of multiple sources, and research project. Similarly, Currie (1993:112) analyzed the conceptual activities demanded in order to write for one introductory university course in business: finding and recording information, using a concept to find and report observational details, using a concept to analyze data, classifying according to a concept, comparing/contrasting, determining causal relationships, resolving an issue, speculating.

This study uses as a resource the ICU task analysis "A Survey on the Use of English in ICU's Academic Programs" (McCagg et al. 1991), which identified such academic tasks as critical reading of academic material, listening to lectures, writing short answer and essay exams, writing essays, small and whole class discussion. These tasks identified by the target academic community itself should be integrated with the students' perceived learning needs in order to create a learner-centered curriculum with tasks appropriate for their academic and long-term language needs.

Summary of ELP Learner Analysis

From the learner analysis which we administered to two freshman classes in the ELP, we identified the students' priorities for studying English to see if their perceived needs were integrated with the teaching agenda of the ELP and the undergraduate program as a whole. In analyzing the data obtained from the survey, we took the mean for each question choice, which gave us the ranking of choice by students, and determined the standard deviation to see the agreement within each question (See Table 1).

Table 1

Q	Mean	SD	Q	Mean	SD	Q	Mean	SD
1b	2.67	1.095	**2b**	2.22	1.098	**3b**	2.89	1.063
1c	3.11	0.919	**2c**	1.83	0.941	**3c**	2.36	0.99
1d	2.67	1.014	**2d**	3.44	1.054	**3d**	1.44	0.877
4a	3.33	0.926	**5a**	1.75	0.874	**6a**	2.06	0.955
4b	1.69	0.856	**5b**	3.69	0.786	**6b**	2.72	1.279
4c	2.22	0.866	**5c**	2.47	0.736	**6c**	2.78	1.333
4d	2.75	1.131	**5d**	2.06	0.955	**6d**	2.17	0.971
7a	3.22	0.898	**8a**	1.72	1.059	**9a**	2.53	1.082
7b	2.14	1.018	**8b**	2.94	1.241	**9b**	2.33	0.956
7c	1.94	0.893	**8c**	1.64	0.798	**9c**	1.69	0.856
7d	2.69	1.215	**8d**	3.14	0.99	**9d**	3.44	0.843
10a	1.86	0.867	**11a**	3.28	0.741	**12a**	2.44	1.107
10b	1.64	0.798	**11b**	1.22	0.485	**12b**	3.42	1.105
10c	3.53	0.91	**11c**	2.86	0.762	**12c**	1.78	0.959
10d	2.94	0.674	**11d**	2.64	1.15	**12d**	1.81	1.009

The following is a summary of several questions chosen from the learner analysis (Appendix B).

Question 1. It is clear from the students' first choice that they are most interested in studying English for the purpose of learning about global issues. Students second choice was tied between learning about American culture and learning travel functions. It is interesting to note that improving grammar is of least importance to them even though it has been the main thrust of their English curriculum through elementary and high school. This question seems to indicate that a global issues curriculum does in fact fit our students' needs.

Question 2. Mean scores for question 2 show a clear ranking for this question. Students chose interpersonal communication as their first purpose for studying English followed by further academic study, job and finally travel. This supports the view that our students need a communicative language syllabus with a focus on academic skills and that they are concerned about their academic study, which is not always the case at universities in Japan.

Question 3. The ranking of this question was at odds with the academic tasks required by the faculty in the ICU survey on the use of English (McCagg et al. 1991). Students deemed speaking most important while reading was least important. Professors on the other hand, valued reading as the most important skill to succeed in their subject, followed by listening to lectures and note taking. Some small classes however, do use discussion and oral presentation. Our language program does meet the students' expectations in that we fully integrate discussion into all tasks and also have one speaking class per week which focuses on communicative strategies.

Question 4. This question also shows clear ranking. Mean scores tell us that small group discussion in English is favored over reading and writing, listening to the teacher and studying grammar and vocabulary as a way of developing language skills. The fact remains however, that ICU faculty regard reading and listening skills as most important. Certainly these skills may be developed within integrated skills tasks which involve discussion. This student preference indicates a general trend towards learner-centered, communicative language education.

Question 8. This question also illustrates the move away from traditional purposes of Japanese education. Students ranked ability to work cross culturally as most important, followed by problem solving, specific training for a Japanese company, and finally the ability to follow instructions. This ranking shows us that according to students, an education which offers opportunities to develop problem solving and cross-cultural knowledge is most meaningful to their real life needs. This fits in well with overall task-based language instruction within a global context.

Question 9. It is well known that the Japanese have been driven to improve their economy since the end of W.W. II, and this is reflected in their educational system as well as the very social and commercial fabric of their society. This question gave us insight into what our students see as being the most urgent issues. Surprisingly, we found that students were most concerned about pollution, followed by global hunger and the threat of nuclear war, and then finally, the economic recession.

Question 10. This question revealed more food for thought. Students ranked newspapers, followed by TV, then school, and finally parents as their main source of information about world issues. A couple of points emerge from these findings. First, there are clearly some discrepancies between what both the professors and ELP staff see as valuable reading material, and what students perceive. Newspaper articles do provide a source of information for some aspects of their academic career, but academic books, scholarly journal articles and textbooks provided a more important source in the eyes of their teachers. Second, if newspapers and TV are the primary sources of information for young people then students need skills in media literacy, detecting bias, and evaluating evidence. This poses the question anew: how should formal education prepare young people to face the 21st century?

Question 11. As global education was originally aimed at geopolitical awareness of educating students for economic means we designed this question in order to find out what our students' vision of global education really is. Students ranked cooperative solutions to global problems as being most important, followed by contributing to a more equal and just world, preparing for a job and finally to help Japan maintain its economic power.

Question 12. Finally, in thinking about preparing for the 21st century, our students ranked flexibility as being most necessary, followed by independence, cooperation and competition. Again, these values show a shift away from the traditional purpose of global education and perhaps even the Japanese way of thinking. Today's global agenda would include flexibility but would place cooperation ahead of independence and competition.

In summarizing the students' perceived needs, we found that they were closely related to both the professors' and the ELP staff's goals and objectives for the undergraduate courses offered at ICU. It appears that students see the content-based curriculum of global education as being viable for their immediate academic goals, as well as compatible with their vision of the future.

Moving From Needs to Task Syllabus

Both professors and students identified the following academic tasks: write a research paper, critically evaluate written text, and make an oral presentation. Students prioritized small group discussion as the most effective way to develop their most highly desired language skills: speaking and listening. From the learner analysis, we added media analysis as an important target task in terms of global

Figure 1

Long's Model for Needs Identification and Task Syllabus (adapted from Michael Long (1985, p. 91)

Target Tasks →	**Task Types** →
Analyse the role of NGO's in Japan	Write a Research Paper
Are Jeremy Rifkin's claims in Chapter One of Entropy justified?	Critically evaluate academic text, scholarly journal articles
Orally present the results of your survey	Make an oral presentation to the whole class
Does " A Fundamental Reassessment of Whaling Activities is Needed" by Fujiwara Eiji, Japan Environment Monitor, present a fair view of whaling in Japan?	Media Analysis
Rifkin claims that the way to solve environmental problems is to consume less. Discuss this solution in small groups	Small group discussion

Pedagogic Subtasks →	**Task Syllabus**
Identify area of research interest, with research questions, locate library resources on topic, gather informtion, define research thesis, submit first draft (summarizing, paraphrasing, citing, synthesizing), second draft (revising organization and coherence), final draft (correcting errors)	Students choose three news articles over the semester about environment. For each article, they write a journal response, to be evaluated on how deeply they reflected on the article and how fluently they communicated their response ↓
Skimming and scanning to focus reading time, using headings, read to locate thesis, topic sentences, and supporting evidence detecting bias	Oral presentation. Students choose one or all of the articles to summarize orally for small group presentation. Elicits brainstorming, provides feedback for research topic.
Organize information with intro middle and conclusion, use rhetorical markers to keep listener's attention, use notecards, use visual props if possible	↓
Read to locate topic and point of view, assess supporting evidence for author's position, check for bias, summarize point of view and evidence; express agreement. Pose open-ended questions which engage group members; summarize others' point of view to clarify; express disagreement politely with good reason, synthesize group member's comments for consensus	Library tour with worksheets on finding sources. Hand in research topic, tentative thesis list of four English sources. Bring first draft of research paper to class for peer editing. Receive teacher feedback on second draft. Hand in final draft.

BEST COPY AVAILABLE

education. The present syllabus for the Environment Unit in the ELP indeed includes activities which match these tasks: students read and discuss three readings on the environment, listen to program-wide lectures around environmental topics, and complete a research paper. However, we feel that by sequencing and integrating these tasks more tightly, content is made more meaningful and therefore comprehensible, and the inter-relatedness of themes, a principle of global education, is demonstrated tangibly for the students. An "Environment" Unit becomes a "Global Education" unit.

The following semester-long project is an integrated skills, task-based activity for Global Education, which prepares students for academic tasks which they will need to do in upper-level ICU classes. Figure 1 shows the move from needs identification to task syllabus. The four steps in Long's model - identification of target tasks, task types, pedagogic tasks, and finally the task syllabus - are represented in four columns. For example, the first item of column one is the specific target task, "Analyze the role of NGO's in Japan." Its more general task type is "Write a Research Paper."

The pedagogic subtasks are several: identify areas of research interest, write research questions, locate library resources on topic, etc. After identifying the task types through the needs analyses, the pedagogic tasks are sequenced in a task syllabus. The last column of Figure 1 shows the sequence of pedagogic sub-tasks which form the core of the global education syllabus: read several news articles about the environment, choose three to respond to in a "response journal" style, make a small group presentation about their articles and generate feedback from their group about possible research topics related to the news articles, write a research paper (three drafts). This integrated skills, task-based activity prepares students for academic tasks which they will need to do in upper-level ICU classes.

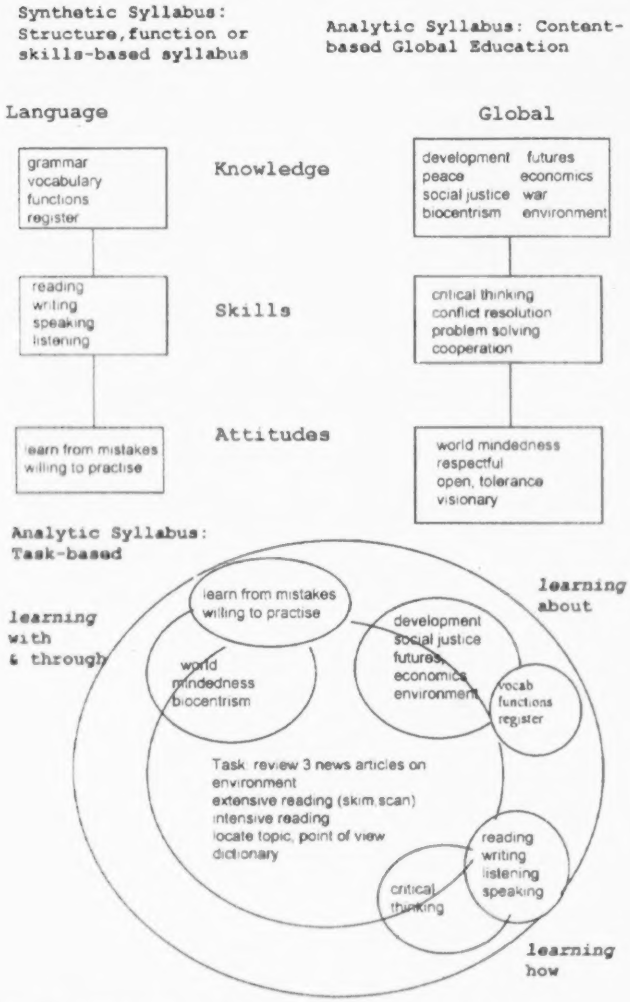
The attempt to integrate synthetic and analytic syllabuses in a task-based syllabus is represented in Figure 2, "Global Education in Task-based Language Teaching." In the upper left corner, a traditional synthetic syllabus can be seen, in which units of language input are grammar, vocabulary, etc. On the right is the content-based analytic syllabus in which there are no discrete units of language input, but rather, themes, topics and issues.

The integrated representation, the task-based syllabus, is shown at the bottom. At the core of syllabus is the task, which in the case of this particular syllabus, begins with extensive reading and reviewing of several news articles, and ends with the academic research paper. Language is acquired as necessary to complete each sub-task, and inter-connected global themes emerge as students continue their research.

As illustrated in Figure 2, students have the opportunity of examining various topics in global education, many of these being interconnected. By reading and writing about interconnected themes, a growing schema builds, facilitating comprehension, and providing for meaningful academic tasks across all language

Figure 2

Global Education in Task-Based Language Teaching



BEST COPY AVAILABLE

skill areas. Examples of research paper topics which students chose are: green consumerism, green economics, steady state economics, waste disposal, river pollution, women's role in the environment, overpopulation, indigenous peoples and environment, Marxism and Capitalism, nuclear waste etc. This list of topics shows how varied and interconnected global education can be. Appendix C contains some excerpts from students' essays and journal responses which show the kinds of topics they chose and how they used the skills, knowledge and attitudes learned from their global education curriculum.

Discussion

There were several problems with the learner needs analysis. A pilot survey would have clarified potential areas of ambiguity in the questions, one being that of time reference with regard to students' needs. Many of our questions did not clearly differentiate between students' perceived language needs for present study in the ELP, future study in ICU, and post-study (e.g., job). For example, question 6 did not specify the context of the writing tasks:

> The types of writing tasks I am most often asked to do are:
>
> ____ to summarize readings
> ____ to research
> ____ to translate
> ____ to analyze cause and effect.

There were no significant tendencies in the group ratings for this question, perhaps because of the lack of clarity in the question. It was also not clear whether students perceived listening and speaking as the most important skills for their present or future academic study, or for their personal life (relationships, jobs, etc.). The task analysis (McCagg et al. 1991), on the other hand, clearly identified academic tasks which are required in the upper level classes of ICU

Assessing task difficulty remains problematic, as Long (1992:46) warns, and deciding where one task ends and the next begins. How many "subtasks" a task type should be broken down into depends on the prior knowledge, ability and skills of the learner, which can be assessed partially by a needs analysis, but which also requires some sort of pre-test if one were setting up a brand new curriculum with unknown students. Our students study paraphrasing and summarizing in their first semester, so didn't require explicit instruction at that level of subtask to complete the tasks required in the Environment Unit of the second semester.

Finally, a more philosophical problem is the place of the learner analysis itself. As Nunan points out, many educators feel that a syllabus should be rooted in the input of experts, not the learners (1988:21). He cites an example from Willing

(1988:78) where students expressed a preference for rote pronunciation practice and total error feedback to pair work and communicative language games. What does the syllabus designer do with the learner's needs analysis when the learner indicates preferred learning styles which are diametrically opposed to current pedagogical practice? Are learners a good source of information for curriculum design? To integrate the learner analysis and the task analysis can be difficult. In our own survey, there was fortunately a great overlap between the two. One area of disagreement was the learners' perceived need for speaking, followed by listening, as the most important language skills for academic success. The McCagg survey indicated that reading was most important. This discrepancy could be a point of discussion in the initial meeting with the students, when the class syllabus is distributed. Fundamentally, if it is possible for syllabus designers to keep in mind the students' larger post-university needs, then the gap between the ivory tower and real life might be lessened. This, of course, is one of the primary goals of global education: to empower students to interact more critically, compassionately, and flexibly with the world they greet upon graduation.

References

Candlin, C. & Murphy, D. (eds.). (1987). *Language learning tasks.* Englewood Cliffs, NJ: Prentice Hall.

Currie, P. (1993). Entering a disciplinary community: Conceptual activities required to write for one introductory university course. *Journal of Second Language Writing,* 2(2):101-117.

Horowitz, D. M. (1986d). What professors actually require: Academic tasks for the ESL classroom. *TESOL Quarterly,* 20(3).

Long, M. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & N. Pienemann (eds.), *Modelling and assessing second language acquisition,* pp. 77-97. Clevedon, Avon: Multilingual Matters.

Lyons, T. (1992). Education for a global perspective. *Orbit,* 23(1):10-12.

McCagg, P. et al. (1991). A survey on the use of English in ICU's academic programs: Implications for the ELP and the CLA. *ICU Language Research Bulletin,* 6(1):48-74.

Mohan, B. (1986). *Language and content.* Reading, MA: Addison-Wesley.

Nunan, D. (1988). *Syllabus design.* Oxford: Oxford University Press.

Nunan, D. (1989a). *Designing tasks for the communicative classroom.* Cambridge: Cambridge University Press.

Nunan, D. (1989b). *Understanding language classrooms.* Hertfordshire: Prentice Hall International.

Prabhu, N.S. (1987). *Second language pedagogy: A perspective.* Oxford: Oxford University Press.

Roald, J. et al. (1990). The global outlook of the class of 1990. Project International Educational Survey, School of Education, Dalhousie and IEC, St. Mary's, Halifax.

Willing, K. (1988). *Learning styles in adult migrant education.* Adelaide: National Curriculum Resource Centre.

Other Recommended Reading

Task-based Language Learning

Long, M. & Crooks, G. (1992). Three approaches to task-based syllabus Design. *TESOL Quarterly,* 26(l).

MacKay, R. (1978). Identifying the nature of the learner's needs. In. R. MacKay & A. Mountford (eds.), *English for specific purposes,* pp. 21-42. Harlow: Longman.

Swales, J.M. (1990). *The concept of task in genre analysis; English in academic and research settings.* Cambridge: Cambridge University Press.

Global Education

Greig, S., Pike, G., & Selby, D. (1987). *Earthrights: Education as if the planet really mattered.* London: Kogan Page.

Hanvey, R.G. (1976). An attainable global perspective. *Theory into Practice,* 21(3).

Pike, G. & Selby, D. (1988). *Global teacher, global learner.* York: University of York.

Reardon, B. (ed.), (1988). *Educating for global responsibility: Teacher designed curricula for peace education, K-12.* New York: Teacher's College Press.

Sokolik, M.E. (1993). *Global views: Reading about world issues.* Boston, MA: Heinle and Heinle.

JALT: *The language teacher*: Special issue on Global Education, May 1993, 17:5.

Content-based Language Instruction

Adamson, H.D. (1993). *Academic competence theory and classroom practice: Preparing ESL students for content courses.* Harlow: Longman.

Spack, R. (1988). Initiating ESL Students into the academic discourse community: How far should we go? *TESOL Quarterly* 22(1):29-51.

Appendix A.

Excerpt from "Questionnaire on the Use of English in Your ICU Class" (McCagg et al., 1991).

Section II: About Reading

1. What percentage of the required reading in your courses is in English?

   A) larger classes _________%
   B) smaller classes _________ %

2. About how many pages of English reading are students expected to do per week?

   A) larger classes: ________ pages per week
   B) Smaller classes: _______ pages per week

3. Please indicate which of the following types of reading are required to be done in English, and indicate what percentage of the required reading each type of text typically accounts for.

	A) Larger classes/ % of reading	B) Smaller classes %
scholarly journal articles	Yes No ___%	Yes No ___%
academic books/chapters	Yes No ___%	Yes No ___%
textbooks	Yes No ___%	Yes No ___%
fiction/poetry	Yes No ___%	Yes No ___%
newspapers	Yes No ___%	Yes No ___%
lecture notes/handouts	Yes No ___%	Yes No ___%
other	Yes No ___%	Yes No ___%

Appendix B.

Questionnaire on Language Learning and Global Education

Rank the following from one (1) to four (4) in order of preference.

1 = most prefer   4 = least prefer

Part A. Language Learning

1. When I study English I want to:
___ learn about world issues
___ learn how to ask for directions, catch a bus and other travel functions
___ improve my grammar
___ learn about American culture

2. My purpose in learning English is for:
___ job
___ further academic study
___ interpersonal communication
___ travel

3. To achieve my purpose in learning English I need to improve my:
___ reading
___ writing
___ listening
___ speaking

4. I believe the most effective way of developing my language skills is:
___ studying grammar and vocabulary
___ small group discussions in English
___ reading and writing in English
___ listening to the teacher

5. In order to succeed in my present university program I need to be able to:
___ discuss in English
___ take multiple choice exams
___ read long texts in English
___ complete an independent research project

6. The types of writing tasks I am most often asked to do:
___ to summarize readings
___ to research
___ to translate
___ to analyze cause and effect

Part B. Global Education

7. Learning about world issues in English will help me:
___ to secure a good job
___ to improve my English
___ to communicate with foreigners I meet
___ to contribute to a more just society

8. I need to develop the following skills for my future career.
___ problem solving
___ specific training for a Japanese company
___ ability to work cross culturally
___ ability to follow instructions

9. The global problems which concern me the most are:
___ threat of nuclear war
___ global hunger
___ pollution
___ economic recession

10. I have received the most information about these problems from:
___ TV
___ newspaper
___ parents
___ school

11. By studying global education I can:
___ help Japan maintain its economic power
___ help to develop cooperative solutions to global problems
___ prepare for a job in an international company
___ help to contribute to a more equal and just world

12. To be an effective citizen in the 21st century I will need to be:
___ cooperative
___ competitive
___ flexible
___ independent

Appendix C.

Students' Writing Samples on Global Issues

Example One

This student's essay deals directly with environmental education and what must be done to heighten the Japanese awareness towards the environment through environmental education.

> Japanese traditional education which emphasizes the amount of knowledge in each subject, has failed to produce a person who is able to perform "systems thinking" or "problem solving" especially in the environment so some change in the curriculum, materials, school systems and teaching strategies will be necessary. More description about environmental problems caused by Japan in the materials would be helpful for students to understand what the reality is, even if it is shocking for them at first. Learners need to know their environmental problems at hand such as over consumption, and make progress to understand how various environment problems and their daily life relate each other. This procedure can be acquired from their real experience as well such as recycling. Also, it is important that this change is brought by not only in education but also by all kinds of positions of Japanese society. Big campaign will be necessary in various work places.

We feel this student has successfully integrated her sources and has posed original and creative solutions at an academic level as well as through real life experiences.

Example Two

In response to a newspaper article on the European and American environmental movement, here is one student's analysis of the problem of environmentalism in Japan.

> The main reason that Japanese politicians doesn't seem reluctant to go to the environmentalism is the close relationship between political and business circles. After W.W. II Japan were desperate to rebuild the destroyed economy and believed that the only way to do was to develop the industry through the cooperation of political and business circles. At present Japan is one of major economic powers, on the other hand, has very much environmental problems caused by radical growth. Japanese begin to find that they need political improvement about environmental issues, but the business circles block the environmental stream by the big power with political circles because it will need enormous money and be unprofitable for business circles to limit the industrial wastes or reduce the amount of

> the use or resources such as tree, oil etc. However, as we know, the environmentalism are getting more and more serious all over the world and if we continue to maintain the present industrial system, the earth will be the place we cannot live. It is necessary for political circles to separate the business circles and enforce environmental policy without intervention of business circles.

This excerpt illustrates how students can consolidate information and interconnect related topics. It appears that this student has used critical thinking and systems thinking in responding to the news article.

Example Three

This last example shows an attitude of world mindedness and illustrates a sense of vision for the future in the solution to "The Effects of the Airline Industry on the Environment." The student took three news articles dealing with environmental pollution caused by airplanes around the world and integrated the points to support his thesis that: "Airline companies should try to make their operation effective by joining hands with other airlines, not only from an economical point of view, but also from an environmental point of view."

(Points of support taken from the essay)

> First of all, aircraft are one of the main causes of air pollution and global warming. Reducing the number of flights by combining flights with other airlines, when they are not crowded, and using airports more efficiently will make this condition better.
>
> Secondly, the noise of airplanes affects the life of people who live near airports. In this case too, airline companies should combine their flights when they are not fully booked, especially at night and early morning.
>
> Another problem concerning the competition of airlines is the waste of resources. The increase of flights and serious traffic jams at airports or in the sky cause a waste of fuel. `Furthermore, there is another waste of resources that must not be overlooked. Of course for the safety, but mainly for their commercial appeal, major airlines frequently change their aircraft. Airline companies, especially the major airlines should maintain their fleets much longer, and should try to recycle retired aircraft as much as possible to save resources."

The examples given suggest that a task-based syllabus within the framework of global education can enhance students' academic skills as well as prepare them with real world skills necessary for citizenship in the 21st century.

FROM TASK DESCRIPTION TO TASK ENACTMENT: TEACHERS' INTERPRETATION OF LANGUAGE LEARNING TASKS

Gertrude Tinker Sachs, Stella Kong, Anne Lo and Tom Lee

Introduction

The Hong Kong government has embarked upon a curriculum reform initiative that will challenge traditional notions of teaching and learning. The Targets Oriented Curriculum (TOC), which was formerly called Targets and Targets-related Assessment (TTRA), is designed to provide a framework for guiding education and ultimately "to improve student learning" (Clark 1993:1). Task-based learning is central to the thinking behind the initiative and according to the Education Department (1992:20) tasks are said to be "the purposeful and contextualized means through which students progress towards the learning targets." Proponents of task-centred or task-based language teaching (TBLT) contend that tasks create conditions for encouraging interactive learning (e.g. Littlewood 1981, 1993, Candlin and Murphy 1987, Nunan 1989, Long and Crookes 1992).

As Hong Kong teachers prepare for TOC's forthcoming implementation, there has been much debate within the education community about teachers' preparedness for curriculum reform. Discussions, both formal and informal, have tended to revolve around (a) questions related to teachers' attitudes toward change, (b) teachers' expertise in terms of knowledge and skills for implementing TOC, and (c) teachers' inherent beliefs about teaching and learning. Related to all these questions are issues of resourcing TOC's implementation, providing ample support for teachers and educating parents and the community at large about the curriculum reform initiative.

The researchers in this project were all seconded to the TOC Teacher Education Team from September 1992 to July 1993. This team was responsible for organising the 3-day introductory seminars to TOC and during this time more than three thousand (3,000) upper primary teachers of English attended courses. They represented the broad spectrum of schools in Hong Kong. It was this exposure to the teachers' fears and perceived limitations in carrying out the assigned teaching tasks that sparked the present investigation. Some problems became apparent during the lesson planning session which required teachers to work with task descriptions. An analysis of teachers' worksheets by one of the researchers indicated that teachers were experiencing some difficulties in designing graded tasks (see Tinker Sachs 1993). However, it was felt that some of these manifested difficulties were also related to teachers' interpretations of the task descriptions with which they worked. These task descriptions were taken from the list of tasks as described in the Education Department's Programme of Study for Key Stage 2 (POS-KS2). Further investigation was necessary to ascertain if our perceptions were ill-founded and premature.

Task descriptions, as they are presently designed, tend to be written in brief, generic terms to invite adaptation for all groups of learners of a particular key stage. Teaching instructions and suggestions do not accompany these descriptions and teachers are therefore called upon to more fully develop these descriptions to accommodate their learners and teaching plans. This implies a certain degree of autonomy and decision making in terms of adapting a task to a particular learning situation. As second language teachers, decisions must be made about the content to be taught and the most appropriate procedures for enacting them. Embedded in these task developments are decisions which are probably influenced by many factors such as teachers' feelings of efficacy, teaching styles, experience, knowledge of learners' proficiency and teachers' skills and knowledge for developing the procedures.

Freeman and Richards (1993), in reviewing the different conceptions of teaching that underlie second language teaching, have called for a movement away from examinations of behaviours and activities of teaching toward an examination of the reasoning and thinking that would motivate teacher behaviours. Earlier, Shavelson and Stern (1981) had identified a general need for more research that described teachers' pedagogical thoughts, judgements and decisions when planning for instruction.

This study is an attempt to examine more closely some of those underlying behaviours that affect and influence the decisions that teachers make when planning for teaching as they work with a given task description. Its purpose is to contribute to the developing body of research on Hong Kong's teachers by focusing on the following questions:

1. What do teachers understand about language learning tasks and what decisions do they make when preparing a task for teaching?

2. To what extent do primary and secondary teachers differ in their interpretations of tasks?

3. What factors influence and affect teachers' interpretations of tasks?

In this paper we will provide some background to the situational context of the investigation, outline our procedures and summarise and discuss the pertinent findings of our investigation. Recommendations will then be made for follow-up research and suggestions made for: (a) teachers, (b) curriculum developers, and (c) teacher educators.

1. Related Studies

Research reporting on aspects of curriculum reform needs to take into consideration some of the related factors which could influence performance and behaviour such as the attitudes and conceptualizations of teachers, background factors such as their professional training and years of experience in the particular field of teaching. Many theorists and researchers have studied the impact of these varying factors. Clark (1988), for example, described certain personality characteristics that predispose a teacher toward change. Fullan (1991) commented on factors such as previous exposure and stage of career.

Recently, a number of studies have begun to provide an understanding of some of the thinking of Hong Kong's teachers. These descriptive studies investigated different aspects of Chinese teachers of English as a second/foreign language such as their attitudes and thinking toward English language teaching. The participants usually came from university and in-service courses and the questionnaire format was the main instrument for data collection.

Young and Lee (1987) argued for a culturally appropriate curriculum in their study. They administered a questionnaire adapted by Falvey (1983) to their primary and secondary teachers to investigate the transmission or interpretation attitude type in teachers before and after a course. Since Young and Lee found a more transmission type of attitude in teachers, they call for a curriculum that recognises these attitudes without trying to change them.

Chan and Lau (1989) found that primary teachers favoured the communicative approach to teaching but found it difficult to apply to the classroom context. Hirvela and Law (1991) had their subjects, primary and secondary teachers, rank their perceptions of their English ability. These researchers found that the majority of participants rated themselves as very good to acceptable while confidence in their teaching ability was rated as strongest in reading and listening with less confidence in their ability to teach writing and speaking. Hirvela and Law also found that teachers who were university graduates tended to have higher self-esteem and self-confidence than teachers who were non-university graduates.

Richards, Tung, and Ng (1992) conducted a more comprehensive survey of the beliefs and attitudes of secondary teachers and found that experienced and trained teachers with recognized teacher qualifications differed significantly from untrained and inexperienced teachers in several respects. One area of difference was in classroom practice activities. Untrained teachers indicated that they required students to memorise dialogues and favoured the writing mode whereas trained teachers exercised greater variety in their selection of teaching activities such as using audio and video tapes. Richards et al. also found a positive relationship between teachers' values, goals and beliefs and teachers' training and years of experience.

More recently, Wan (1993) investigated secondary school teachers' response to issues related to teacher participation in curriculum development. Wan's results indicated that while teachers were able to design their own activities for students and facilitate student learning by using new techniques, very little team teaching and collaboration to introduce change took place.

A final study of relevance here was conducted by Carless and Lee (1993). They found that participants on a 16-week in-service course were more receptive toward communicative language teaching and TOC after the course than before the course.

While the focus of these studies and their results have been varied, put together, they provide an emergent picture of Hong Kong's teaching practice. Bearing in mind the constraints of generalising from the limited number of studies reported here, there appears to be a noticeable difference in performance and attitude based on years of experience and training. Bickley (1987) noted that teachers needed to be prepared for responsibility and autonomy in their teaching. In this study, we contribute further to the picture of Hong Kong's educators by studying the decisions that teachers make as they plan for task-based instruction.

2. Method

Participants: A total of 114 teachers (43 secondary, and 71 primary) ranging in age from 21 to over 50 years participated in this study. They all took part in a 16-week in-service refresher course at the Institute of Language in Education (ILE) during the first half of the 1993-94 academic year. Background information collected at the beginning of the study showed that approximately one-third of the these teachers were holders of bachelor or higher degrees and twice as many held teachers' certificates. Roughly two-thirds had English as their major or one of their majors in university or at the teachers' training colleges. Roughly one-third had 1 to 5 years teaching experience, another one-third, 6 to 10 years and the remainder 11 or above. And one in seven were either assistant master/mistress (AM), senior graduate master/mistress (SGM), or panel-chair-persons (PCP) in their schools. Therefore, these participants represented a wide range of teaching and administrative experience.

Procedures: Teachers were required to work with a learning task from the TOC Programme of Study - Key Stage 2. For purposes of comparison between the participants, only one task was chosen. The main criterion for selection was degree of "openness". Openness was interpreted as the degree of malleability allowed by the task. This task feature facilitated adaptability for learners of different levels within grades and across the grade range of upper primary and lower secondary. The task chosen needed to afford teachers numerous possibilities in design, procedures, activities, and selection of teaching content. Also of importance were interest and appeal to learners. Based on the aforementioned criteria, a task called *Watching T.V.* was chosen (see Appendix).

Administration of the Lesson Planning Sessions: The lesson planning sessions for the secondary and the primary teachers were conducted in 2 separate cohorts in September at the beginning of the course. The time was chosen to avoid contamination from course work on lesson planning or task-based learning. To reduce the possibility of the Hawthorne effect, teachers were not informed of the research purposes of the assignment until after the collection of the lesson plans.

Instructions to Teachers: Teachers sat in one large classroom to design their lesson plans. They were instructed to develop lesson plans for use in the teaching units on *Task* and *Lesson Planning* which were forthcoming. To encourage the development of "good" lesson plans, teachers were asked to design the lessons for an inspector's visit to the class. Teachers were given one hour to do the work. No guidelines were given for designing the lesson plans but teachers were directed to develop a lesson plan that was appropriate and applicable to the needs and interests of the group of learners they had specified.

Post-Questions: Following the design of the lesson plans, participants were asked to respond to six post-questions. These were set to capture the teachers' thoughts on the work they had just completed (see Table 1).

Interviews: Following the collection of data, informal interviews were conducted with several of the participants. This was to obtain additional insights on the lesson plans the teachers had developed and any methodological problems they might have encountered in doing them.

Criteria Development and Rating of the Lesson Plans and Post-questions: The first phase of our criteria development involved impressionistic marking of twenty percent of the lesson plans by the researchers. After extensive discussion, the criteria for analyzing the lesson plans were agreed upon (see Tables 2 and 3). The two categories established were (a) *Task Interpretation* which would capture whether teachers could translate the task into learners' needs by applying their knowledge and experience, and (b) *Holistic Interpretation* which would capture how liberal teachers were with the task in terms of adapting the procedures for their purposes. For the post-questions, categories were created for classifying the responses (see Table 3).

Once the criteria for grading the lesson plans and post-questions were agreed upon, the inter-rater reliability was computed by randomly selecting 20% of the lesson plans and post-questions to be graded by the four raters. Inter-rater reliability was generally high, ranging from .7 to 1 (indicating perfect agreement among raters). The complete inter-rater reliability results are reported in Appendix Table 1.

3. Results and Discussion

First, the findings of interest from the questionnaire will be reported followed by the results from the lesson plans and post-questions. Next, some interesting relationships between the different variables will be reported.

Table 1

Response Categories For Post-Questions

Question Number	Categories
1. Which part of the lesson is the most important part for pupils' learning?	1. Language 2. Task / Activities 3. Language and Task / Activities 4. Others
2. What was your initial reaction to the task?	1. Teachers' Constraints 2. Teachers' Positive Aspects 3. Students' Constraints 4. Students' Positive Aspects 5. Specific Task Constraints 6. Task Strengths 7. Mixed Response 8. Others
3. What difficulties did you have in the process of planning the lesson?	1. Classroom Management and Procedures 2. Objectives 3. Language Use 4. Resources 5. Time 6. None 7. Others
4. Would you teach this lesson to your class? Why or why not?	1. Yes 2. No 3. Maybe / Unsure
5. What will pupils learn in this lesson?	1. Form 2. Function 3. Affect 4. Form and Function 5. Form and Affect 6. Function and Affect 7. Form, Function and Affect 8. Nothing
6. Is this lesson similar to the lessons you presently teach?	1. Yes 2. No 3. Unsure

Table 2

Criteria For Task Interpretation

Category	Good	Average	Poor
Interaction	good interaction	some interaction	lack of interaction
Language Support	full language support	adequate language support	inadequate or lack of language support
Development	sound development of procedures	adequate development of procedures	insufficient details, lack of coherence, lack of clarity in the procedures
Application	broad application, higher level thinking encouraged	some application	no application of the task

Questionnaire

The questionnaire used in the study was adapted from the one developed by Richards, Tung, and Ng (1992). The first part of the questionnaire was a brief survey to gather participant background information, and was reported earlier. The second part of the questionnaire was organised around the issue of how teachers feel about teaching and learning English in Hong Kong. An overwhelming majority, 80%, of the teachers indicated that they liked teaching English while only about 6% said they did not. Their students, however, did not seem to show as much enthusiasm, for only about 40% of the teachers thought that their students wanted to learn English, whereas 48% indicated that their students did not. Nevertheless, on the question of how important English is in Hong Kong, opinion was one-sided with 98% indicating "yes" it was important.

The third part of the questionnaire addressed the issue of the degree of autonomy teachers felt they had in their work. The question was "To what extent do you feel you can make decisions about different aspects of teaching?" The results indicated that in some aspects such as "activities done in class" and "sequencing of teaching content", secondary teachers enjoy slightly greater autonomy than their primary school counterparts.

The fourth part of the questionnaire centred on the question of teachers' self-confidence and asked teachers to indicate how much they knew about different topics. The results here showed that a few areas may perhaps raise concern. One is a significantly low percentage of both primary and secondary school teachers who considered themselves not knowing much about how children learn foreign and second languages. Areas such as English language systems and curriculum design for English also seem to imply that primary teachers may experience difficulty in coping with them.

The last section of the questionnaire was a survey of the kinds of activities that teachers use. The results confirm that traditional classroom activities in Hong Kong remain very popular with teachers, primary and secondary alike. Activities such as dictation, reading aloud, memorizing dialogues and grammar exercises were all rated high, whereas drama, songs, choral speaking and projects came low on the teachers' list. In some aspects, primary and secondary teachers differed quite significantly. Secondary teachers seemed to favour memorizing dialogues, pair/group work and composition much more than primary teachers. Meanwhile, there seemed to be some encouraging signs which indicate that the primary and secondary classrooms might be moving towards a more active approach with the increasing adoption of games, role play and group work.

4. Lesson Plans

Task Interpretation

As expected, the distribution of lesson plans for the categories *good*, *average* and *poor* was approximately normal for both the primary and secondary teachers. The majority of the responses were assigned to average -- 57.1% for primary and 45% for secondary. However, slightly more lesson plans were assigned to poor -- 25.7% for primary and 32.5% for secondary -- than to good category -- 17.1% for primary and 22.5% for secondary.

Holistic Interpretation

More primary lesson plans were assigned to the *expanded* category, 31.3%, with the smallest number being assigned to *exact*, 19.4%. For secondary teachers, 53.8% of the lesson plans were categorised as *amended* while 7.7% were categorised as *abridged*.

Task and Holistic Interpretation

There was a significant relationship between *task interpretation* and *holistic interpretation* for primary ($\chi^2 = 15.62$, $\underline{p} < .02$) and secondary ($\chi^2 = 16.76$, $\underline{p} < .01$). For primary teachers this meant that the majority of expanded tasks were good to average. For secondary teachers, the expanded lesson plans also ranged from good to average. Thus, expanded tasks generally received a higher rating. This trend was also observed for the amended tasks while the exact and abridged tasks tended to range from average to poor. This tendency for expanded and amended to get higher ratings is not surprising as the criteria for the good assignment were more fully exploited in amended or expanded while poor seemed to fit the criteria associated with the more limiting categories of exact and abridged.

Table 3

Criteria For Holistic Interpretation

(Category: Key Feature)

Exact	Abridged	Expanded	Amended
No change in the task's order	Omission of one or more of the task's components	Expansion of the task by the addition of one or more components	Different ordering OR minor changes OR edited theme/ content

*Note: Irrelevant and unrelated descriptions were given an "ignored the task" categorisation.

Post-questions

In response to the first question "Which part of the lesson is the most important part for pupils' learning?", both groups of teachers recognised *the task/ activity* as the most important part of the lesson, 81.4% for primary and 69.2% for secondary. Likewise, in question two, "What was your initial reaction to the task?", both groups' reactions recognised *task strengths*, 44.3% for primary and 42.9% for secondary. This was followed by *teacher constraints*, 16.4% for primary and 22.9% for secondary. In the third question, "What difficulties did you have in the process of planning the lesson?", both groups of teachers indicated that *classroom management and procedures* accounted for most difficulties, 42.9% primary and

27.8% secondary. Next was *language use* with 28.6% for primary and 22.2% for secondary. In response to question four, "Would you teach this lesson?", the majority of both primary and secondary teachers said *yes* they would, 68.6% primary and 79.5% secondary. Both groups noted *function* for question five "What will pupils learn?", 50% for primary and 57.5% for secondary. The only difference in the response for both groups was the last post-question, number six, "Is this lesson similar?" Of the primary respondents, 72.3% said *no* while 70% of secondary teachers said *yes*. These results indicate that the majority of teachers recognised the intrinsic teaching value of the task that was chosen. They also acknowledged the task's main components as the major activities of the lesson. However, there appears to be a dichotomy between what primary teachers recognise as valuable and what they actually do in practice. These results suggest that primary and secondary teachers are similar in thinking but not in practice.

Holistic Interpretation and what will pupils learn?

53.8% (*n*=57) of all the teachers combined indicated that *function* was what pupils would learn. These teachers also tended to design lesson plans that were categorised as amended or expanded ($\chi^2 = 24.7$, $\underline{p} < .05$). This trend was also observed for teachers who indicated a combined response (i.e. function and form or function and affect etc.) as these lesson plans were more consistently assigned to the expanded and amended categories. This relationship implies that teachers' awareness of the broader "functional" aspect of language as opposed to the more traditional "form" may be associated with other factors which would have some influence on the type of lesson teachers design.

Attendance at TOC seminar and is this lesson similar?

76.9% of those participants who had attended TOC said that their lesson plans were *not similar* to the way they presently teach ($\chi^2 = 4.9$, $\underline{p} < .03$). Only 23.1% of those who had attended TOC said this lesson was similar. Of those who had not attended the TOC seminar, half, 50.6%, said that the lesson was similar to the way they presently teach. Despite the fact that TOC participants could recognise the strength of a task, they failed to teach in this way despite having attended the three-day introductory TOC seminar. However, it can be said that the TOC seminar may have helped teachers to be aware of how different their lessons are from task-based teaching. Those teachers who had not attended TOC and said that their lessons are similar to the one they designed may be unaware of *how* different their lessons are. Experience tells us that the majority of primary teachers **and** secondary teachers do not generally practice task-based/ communicative type teaching. While teachers may have the desire to teach this way, more support at all levels is needed to make it possible for them to do so.

Number of Lesson Plans

Of the combined primary and secondary respondents who indicated that they had designed ***no lesson plans*** for a given term, 72.7% said that the lesson was not similar to the way they presently teach ($\chi^2 = 7.7$, $p < .02$). On the other hand, respondents who had designed 1-5 and 6+ lesson plans per term tended to say the lesson was similar to the way they presently teach. These results seem to indicate that those teachers who design lesson plans tend to be teachers who do a certain type of teaching, in this case teaching that encourages the use of a task-based curriculum.

Qualifications

Results for primary and secondary teachers indicated that 70.8% of the teachers who had a university degree said the lesson was ***similar*** to the way they presently teach compared to 65.3% of those who had a teacher's certificate and who said the lesson was ***not similar*** ($\chi^2 = 8.2$, $p < .004$). These results signify the possible differences in practice between teachers with different qualifications.

Teachers' Years of Experience and Task

Good to *average* ratings tended to be assigned to combined primary and secondary teachers who had fewer years of experience in teaching English. More experienced teachers got poorer ratings ($\chi^2 = 15.68$, $p < .05$). This unexpected negative correlation may be attributed to the younger teachers' recency of training, exposure to more progressive methods and possible receptiveness and openness toward change. These results are inconsistent with the findings in the literature which commonly finds a positive correlation with years of experience (e.g. Shavelson and Stern 1981, Nunan 1992). While this is good news for novice teachers, more experienced teachers may be exercising restraint in their designs. These teachers may feel that little can be done to change the conditions in their schools, such as the large number of students and small classrooms, and this attitude may be reflected in their lesson plans. More experienced teachers may also have a view of "good" teaching that differs from younger teachers. More experienced teachers may want to change their perceptions by updating their knowledge of teaching to improve not only the design of their lesson plans but ultimately, practice.

Knowledge and decision making

Question number 16 on the questionnaire asked teachers to what extent they could make decisions on a number of items (e.g. how and what to teach and the amount of home work). Question number 17 asked teachers to respond to how much they knew about different topics such as how to teach a second/foreign language or task-based learning. A significant positive correlation of .29 ($p < .01$) was observed between question 16 and ***would you teach this lesson***. Significant positive correlations of .39 ($p < .01$) and .41 ($p < .01$) were also observed between questions 16 and 17 respectively. These observed positive relationships are not surprising because low or high feelings about one's knowledge would affect how one feels about the degree of decision making one can make. These results mean that high or low ratings of decision making and perceptions of knowledge could be an indicator or whether or not one would teach a certain way.

5.Interviews

Interviews with several of the participants indicated that some of them had experienced difficulty in developing their lesson plans as they were unfamiliar with the components of a lesson plan. Their comments were consistent with the answers that teachers had written in response to the post-questions.

6. Limitations

One major constraint with this type of study is that teachers' plans do not necessarily reflect what they would actually do in the classroom context. A further line of investigation would need to pursue the extent to which teachers' pre-planning affects and influences actual classroom procedures. Additionally, the actual designing of the lesson plan itself may have been a limiting factor in this study as many of the respondents indicated in follow-up interviews that they had not done a lesson plan before and did not know what to include or how to write it. Their performance may also have been influenced by a lack of information such as objectives and inadequate information of the television programmes. Even though a list of television programmes was provided for teachers and may have itself influenced their lesson designs, teachers' limited knowledge of the programmes may have been an added constraint. A final limitation may be related to the task that was chosen for the teachers to do. Only one task was assigned in this study and this limits the extent to which generalizations can be made about teachers' interpretations of other types of tasks.

7. Conclusion

In this study we have described the interpretations that teachers made when they were asked to develop teaching plans for a given task. The indication from this research is that teachers do have a tacit understanding of how to develop and adapt a task description. Most teachers included an element of interaction in their lesson plans and tried to provide some language support for students. They recognised the larger purpose of the task by noting that function was what students would learn. Teachers also exercised judgement when they made decisions to change the procedures and content of the task. While all this is encouraging, curriculum developers and TOC teacher educators may want to provide more seminars and resources for teachers in developing their classroom management and procedures along with providing examples of how teachers can give adequate language support

BEST COPY AVAILABLE

to pupils. These areas were the major sources of difficulties identified by teachers in the development of their lesson plans.

The results also indicate that there is a disparity between what teachers know and desire and what they actually practise. This is particularly true with primary teachers who differ from secondary teachers not only in practice but in their self-perceptions on decision making and knowledge. Further investigation is needed to identify the extent to which teachers' conceptions and behaviours are internally or externally guided. Additionally, there is a need for the general upgrading and promotion of teachers' qualifications and feelings of efficacy within the profession.

As a result of this study, we want to explore further issues related to the viability of lesson plans in task-based learning. Nunan (1992:136-138) notes that studies on interactive teaching are rare and that there are only a few on language teaching. We would want to take up this challenge and examine the extent to which primary and secondary, and experienced and inexperienced teachers differ in their planning and interactive enactment of those plans in task-based learning situations.

Acknowledgement

We gratefully acknowledge the support of lecturers from the ILE for assisting with the development of our criteria, Hon Bon Lee for data entry, John Sachs for statistical analyses and Peter Tung and Raymond Lam for input on the design of the study. Our special thanks to one anonymous reviewer for comments on a draft of this paper, and last but not least, all the participants of the ILE 932 courses without whom this study would not have been possible.

References

Bickley, V. (1987). Developments in English language teachers' education. In R. Lord & H.N.L. Cheng (eds.), *Language education in Hong Kong*, pp. 187-217. Hong Kong: Chinese University Press.

Candlin, C. & Murphy, D. (eds.) (1987). *Language learning tasks*. Englewood Cliffs, NJ.: Prentice Hall International.

Carless, D. & Lee, I. (1993). Attitude change on an INSET course for primary teachers of English. Paper presented at the First International Language in Education Conference. University of Hong Kong: Hong Kong.

Chan, A. & Lau, M. (1989). Primary ELT and language teacher education - A Hong Kong perspective. *ILEJ*, 6:72-82.

Clark, J. (1988). The relationship between teacher in-service and educational change. *ILEJ*, 4:30-38.

Clark, J. (1993). The Hong Kong model for TTRA standards design: Origins and major characteristics. *Curriculum Forum*, 3(1):1-5.

Education Department (1992). *Programme of Study for Key Stage 2*. Hong Kong: Government Printer.

Falvey, M. A. (1983). Teacher attitude and teacher behaviour. M.A. dissertation, University of Birmingham.

Freeman, D. & Richards, J.C. (1993). Conceptions of teaching and the education of second language teachers. *TESOL Quarterly*, 27(2):193-216.

Fullan, M. with Stiegelbauer, S. (1991). *The meaning of educational change*. (2nd Edit.), London: Cassell.

Hirvela, A. & Law, E. (1991). A survey of local English teachers' attitudes toward English and ELT. *ILEJ*, 8:25-38.

Littlewood, W. (1981). *Communicative language teaching: An introduction*. Cambridge: Cambridge University Press.

Littlewood, W. (1993). Cognitive principles underlying task-centred foreign language learning. In N. Bird, J. Harris & M. Ingham (eds.), *Language and Content*, pp. 39-55. Hong Kong: Government Printer.

Long, M.H. & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly*, 26(1):27-56.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.

Nunan, D. (1992). Teachers as decision makers. In J. Flowerdew, M. Brock & S. Hsia (eds.), *Perspectives on second language teacher education*. Hong Kong: City Polytechnic of Hong Kong.

Pajares, M.F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3):307-332.

Richards, J.C., Tung, P., & Ng, P. (1992). The culture of the English language teacher: A Hong Kong example. *RELC Journal*, 23(1):81-102.

Shavelson, R.J., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgements, decisions, and behaviour. *Review of Educational Research*, 51(4):455-498.

Tinker Sachs, G. (1993). Graded tasks and learner differentiation in the classroom. *ILEJ*, 10:101-120.

Wan, Chan K.K. (1993). Reflections on teacher participation in curriculum development in Hong Kong. *Curriculum Forum*, 3(3):6-10.

Young, R., & Lee, S. (1987). EFL curriculum innovation and teachers' attitudes. In R. Lord & H.N.L. Cheng (eds.), *Language education in Hong Kong*, pp. 83-97. Hong Kong: Chinese University Press.

Appendix

Lesson Planning Session

Purpose:

This session on lesson planning is preparatory work for the seminars on tasks and lesson planning. Your lesson plans will be useful for these sessions.

Direction:

Your task is to design a lesson plan based on the task description given below. The lesson plan should be appropriate to the needs and interests of your pupils. Design the lesson as you would for the school inspector's visit to your class.

Please work alone for this assignment.

INSTRUCTIONS

Read the task given below carefully and then complete the steps on the next page:

Watching T.V.

Learners, in groups, ask each other which T.V. programmes they like best and which they dislike and why. They then ask each other about the programme which their parents, brothers and sisters like. Then they give a short talk to another group on the most popular programmes for themselves, fathers, mothers, younger/older brothers and sisters.

Table 1

Inter-Rater Reliability

Lesson Plans

Rating Item	Alpha
1. Task Interpretation	0.8
2. Holistic Interpretation	0.7

Post-Questions

Question Item	Alpha
1. Which part of the lesson is the most important part for pupils' learning?	0.9
2. What was your initial reaction to the task description?	0.8
3. What difficulties did you have in the process of planning the lesson?	0.8
4. Would you teach this lesson to your class?	1.0
5. What will pupils learn in this lesson?	0.8
6. Is this lesson similar to the lessons you presently teach?	1.0

PART THREE

TEXT AND TALK IN LEARNING

INTRODUCTION

In exploring the relationship between language and learning, it is essential that both the written language and the spoken language are examined. Hence the sub-theme "text and talk". Under this sub-theme, the following issues were specified for discussion:

1. Issues in using L1/L2 as the medium of learning
2. Classroom talk in learning and language learning
3. Learning to write and writing to learn
4. Text types and language teaching
5. Assessing spoken and written language

This relatively short section consists of four papers, three of which deal with text and one with talk. Though different issues were addressed in these papers, they share certain common features. Firstly the subjects in the studies were all Hong Kong students either in secondary schools or tertiary institutions except for *Falvey and Sengupta's* paper in which some of the data were taken from corporate writing in Hong Kong. Secondly, in most cases; both qualitative and quantitative data were examined. Thirdly, the findings in these papers have very practical implications for classroom teaching.

Revision in writing has often been understood as making modifications to the presentation and development of ideas and the concerns of these modifications are mainly stylistic and organisational. *Falvey and Sengupta* maintain that in fact revisions often result in a change the propositional content and that the modified proposition is better considered as a new proposition rather than a variant of the same proposition. Examples mainly from secondary student writing and some from graduate student writing have been quoted for illustration. They argue that an awareness of propositional modifications is essential in achieving consistency and coherence in writing since the modification of one proposition often necessitates corresponding modifications of subsequent or preceding propositions.

Also examining student writing, *Lewkowicz* looks at the issue of whether providing source materials for a writing task actually hinders or facilitates student writing. Research on writing suggests that providing source materials not only gives students something on which to base their writing but is also a closer approximation of authentic academic writing tasks. Lewkowicz's study of the writings of seventy five non-native English speaking undergraduates, however, shows that this is not necessarily the case. In comparing the writings of the group being given source materials (Group 1) and the group not being given source materials (Group 2), she found that in terms of mean length, there was no significant difference. Moreover, although there were more ideas in Group 1 writings, the ideas in Group 2 writings were more fully developed. There was also plenty of evidence of exact and near copying or plagiarism in Group 1 writings, something which is commonly found among students in Hong Kong. Lewkowicz concludes that providing source materials, contrary to some research findings, may actually hamper students' writing rather than enhance it.

The issue of plagiarism in student writing is a serious one. Sometimes students plagiarise because they feel that the writer can express what they want to say much better than they can, as *Lewkowicz* found in her interviews with the students. Other times, it could be because they cannot quite understand what the writer is saying and yet somehow they feel that the point made is important enough to be included in their writing. In other words, how well students understand the text could be a contributing factor. *Icy Lee's* paper investigates whether signals can help reading comprehension. Specifically, she focused on three types of signals: headings, previews which summarize the main ideas, and logical connectives. These three signals were added to an expository text both individually and in combination, and administered to 285 S4 (Grade 10) students to examine the effect on their reading comprehension. She found that headings and previews helped comprehension, but not logical connectives. She then discussed in detail the possible reasons why adding logical connectives did not aid comprehension, a topic which several papers presented at the conference addressed. The findings of *Icy Lee's* study, as she points out, have practical implications for the design of reading materials for ESL readers not only in the English language subject but also other content subjects.

Finally, *Annie Lee's* paper, which is the only paper dealing with "talk", investigates the vocabulary teaching strategies of ESL teachers in Hong Kong secondary schools and their effects on vocabulary learning. She first recorded and transcribed four lessons and drew up a list of vocabulary teaching strategies used by four ESL teachers in two secondary schools. She found that some teachers used only one strategy and others used a combination of strategies. She then conducted an experimental study on four groups of ESL students (each group of about 20) to investigate the effectiveness of using only one type of explanation (mono-type) and a combination of types (multi-type), as well as the effectiveness of teaching vocabulary items with and without a text. Her findings show that while the groups taught with a text did significantly better than the groups taught without a text, there was no significant difference in the type of explanation given. She concludes that context in vocabulary teaching is very important. She further observes that the lack of difference in explanation type shows that there could be other factors operating such as the nature of the vocabulary item itself, how far the teacher involved the learner in the explanation, and whether an item has been over-explained. Her findings and observations are very useful to classroom teachers who wish to teach vocabulary effectively.

Amy Tsui

HOW STUDENTS REVISE PROPOSITIONS: AN EXPLORATORY STUDY OF PROPOSITIONAL MODIFICATION

Peter Falvey and Sima Sengupta

This paper describes an exploratory study of a linguistic phenomenon which will be referred to, in the course of the paper, as **propositional modification.** The phenomenon was discovered initially during an analysis of the revised texts of corporate writers (Falvey 1993). Having noted this phenomenon, the writers of this article decided to investigate the revised texts of other types of writers such as teachers and students. This article discusses the results of an initial exploration of the texts of Hong Kong secondary school students to ascertain whether the same phenomenon occurred in their texts.

The paper is divided into six parts:

- The first part describes the theoretical underpinning to studies of propositional meaning.
- The second part discusses studies of propositional meaning in previous work and shows how they differ from the phenomenon discussed in this paper.
- The third part briefly charts the process of the discovery.
- The fourth part presents data arising from the analysis of secondary student texts.
- The fifth part discusses pedagogical and research implications arising from the study.
- The final section concludes the article.

Propositional meaning

The word **proposition** is often used synonymously with the word **statement**. Propositions occur as declarative sentences. They do **not** occur in questions, exclamations, or commands. Bell describes a proposition as:

> the unit of meaning which constitutes the subject-matter of a sentence (and, once realized in actual use, that of the utterance as well) (Bell 1991:107)

An earlier definition from Hurford and Heasley is:

> that part of the meaning of the utterance of a declarative sentence which describes some state of affairs (Hurford and Heasley 1990:19).

In creating a proposition a writer/utterer makes an assertion. Each assertion has a truth value. This means that the assertion can be proven or disproved as a matter of fact. However, if the assertion is considered to be not a declaration of a fact but an opinion, that opinion can be challenged or confirmed. For example, the assertion 'It is cold today' is a proposition which can be proven or disproved if the temperature outside is 32 degrees centigrade. Also an assertion such as 'This ice cream is not as creamy as yours' can be agreed with, in which case the utterer's opinion of the truth value of the assertion is confirmed, or it can be disagreed with, in which case the truth value of the assertion is challenged. The challenges to or confirmations of the original proposition can be either those of opinion (what the respondent asserts as a belief based on experience or intuition), or of fact (knowledge of the cream content of each type of ice cream as stated on the contents package).

The standard way of representing a proposition in Logic, from which the notion is derived, is

$$P$$

However, in research literature, rarely is this representation of a proposition seen as a lone representation. It is more often seen as follows:

$$P = Q$$

This means that one proposition **P** equals another proposition **Q**. As an example let us take the following propositions:

1. James likes Mary
2. Mary is liked by James

In this case proposition 1 and proposition 2 can be said to have the same truth value. Thus **P=Q.** The truth value of the two propositions remains the same even though the syntax of each is different. Copi and Cohen (1990:5) state, in reference to propositions such as the two above:

> We use the term "proposition" to refer to what such sentences as these are typically uttered to assert.

In similar ways, propositions can have the same truth value even if the language in which they are uttered changes, for example:

3. It is raining (English)
4. Il pleut (French)

5. Está lloviendo (Spanish)

In each case the truth value of the utterance 'it is raining', though expressed in three different languages, remains the same.

Propositions are central to meaning in discourse. However, they cannot be interpreted as clauses or sentences (Bell 1991:106-127). Linguists often see a proposition as the smallest conceptual unit of communication, as 'the cognitive counterpart of a clause' (Callow and Callow 1992:6). It is through the relationships which are built up between propositions that the coherence in a text or message is established. Frederickson, (cited in Cooper and Greenbaum 1986:227-267) in describing how readers and writers construct 'representations' (p. 227) of the meaning and language structures in text, distinguishes between textual structures by which meanings are encoded and communicated and conceptual structures by which knowledge from texts is represented in memory and subsequently realised as utterances. Conceptual structures, he believes, are represented as propositions and frames while clauses and sentences encode meaning at textual level.

Propositional modification in writing research

Some research in composition writing has focused on the nature of propositional development in which a proposition (or an idea-unit) (Sato 1990) is expanded, elaborated or justified by the addition of extra text to the original proposition. Text linguists too have concerned themselves with the ways in which propositions are developed and elaborated (Mann and Thompson 1986, 1987). Researchers, such as Philipson (1991), have shown that assertions followed by specification (i.e. a proposition followed by relational propositions) improve the coherence of a text.

However, in recent work (see for example, Falvey 1993), it has become evident, in a study of the revision strategies of corporate writers, that propositional modification is an important phenomenon in revision which deserves special attention. By propositional modification we mean:

the change(s) made to a proposition expressed in an original text such that the modified proposition:

(a) no longer contains the same truth value as the original proposition

(b) must be considered as a new proposition.

Falvey's study examines consecutive texts written for the same intended audience by two different writers or by the same writer who, at a later date, revised the original text. In comparing these texts, he notes changes in the communicative intent of the revisers which were made explicit by the modification of propositions in the original text.

In the following example (Falvey 1993:453-454), we can see how the original main proposition of Example 1a (in bold) is modified by the reviser (the original writer's Line Manager) in Example 1b. The original writer, working once more on the text in Example 1c, demonstrates that as far as she is concerned, the modification does not contain the truth value that she considers appropriate to this part of the text and modifies it once again to match the original proposition. In doing so, the original writer also adds an elaboration to the proposition in order to justify the assertion which it contains viz.

Example 1a	Being an independent organisation separate from the government, **every aspect of our work....**
Example 1b	Being an organization independent from most aspects of Government, **much of our work** .....
Example 1c	Being an organization independent from most aspects of Government, and as suggested by Sir Alistair Blair-Kerr, **every aspect of our work** ....

The original writer (Examples 1a and 1c) has made it clear, that for her, the truth value of the proposition asserted in Examples 1a and 1c is that, as a consequence of being largely independent of government, **every** aspect of its work is affected. This assertion stresses the wholeness of the work, not the partial definition which the reviser was attempting to assert in Example 1b. The propositional weakening suggested by the reviser is rejected.

In many cases cited in Falvey, the analysis revealed that the reviser (whether it be the original writer or a new reviser) had decided that the original proposition was either too strong or too weak and that it needed to be either strengthened or weakened in order to match the communicative intent of the reviser. When this occurred, modifications were made to the original proposition which changed it so that meaning in the modified proposition was then construed differently.

A modified proposition can differ from an original proposition minimally in surface morpho-syntactic terms. A change of modal, the addition or deletion of one word can suffice to modify the original proposition. (Further examples from authentic texts are given in the section on student writing.)

The process of discovery

Having noted the phenomenon of propositional modification in corporate text revision, it was felt that an exploratory investigation of whether this phenomenon occurs in revisions performed by students at secondary schools was worthwhile. Such an investigation would reveal whether this phenomenon was associated only with the revisions of mature, proficient, second language writers who were the subjects of Falvey's study or whether it would be apparent in the revised texts of secondary school students.

An examination of the revised texts of secondary school students, who were learning revision strategies as part of an intervention study (Sengupta, ongoing) provided evidence of this phenomenon of propositional modification being used by students, sometimes in response to teacher or peer questions, and at other times as a strategy for making logical connections.

In typical student composition, a proposition is often asserted and then either left as it is alone without any back-up, or it is backed up with justification or exemplified with elaboration. In the section that follows it will be seen that in the revisions of students, they, just like the more mature and proficient, adult, corporate writer/revisers in Falvey's study, are also capable of making strategic revisions which change the truth value of their original proposition through a process of propositional modification in an attempt to match the text more adequately or appropriately to their communicative intent.

Data from student writing

This section presents data from two sources, the texts of Hong Kong secondary school students in S4-5 (Grades 10-11) and the text revisions of a graduate student. In both sources there is evidence of both propositional modification and elaboration in their revisions. The data from Hong Kong secondary school students provides the major source for description in this section. The data from the graduate student is used to demonstrate that the phenomenon occurs across ages.

Secondary school writing

The students in this secondary four classroom were revising their first draft of an article for the school magazine, which was entitled :"Are schools responsible for student suicides?"

Text 1. (addition: in bold in the revised version)

Original: It is undeniable that schools are responsible for student suicides

Revised: It is undeniable that schools are **largely** responsible for student suicides.

NB: The syntactic errors have been deleted.

In text 1 the reviser has added one word, an adverbial, to the original proposition. With this addition the original proposition has changed since in the original proposition the writer was situating the blame for student suicides on the schools alone, while in the revised proposition the writer is only partly blaming the school. Therefore, the truth value of the proposition **P** of the original text 1 is not equal to the truth value of the proposition **Q** in the revised text 1. Hence the writer has modified the original proposition.

When comparing Text 1 with Text 2 below, it is interesting to note that in Text 2 the truth value of the original proposition is the same as in the revised version i.e., 'schools should be blamed.' We can thus say that in the original and revised versions of Text 2 that P=Q. No propositional modification has taken place. The addition of a clause (in bold type) is an elaboration of the original proposition with an explicit relational clause. Here the writer has not changed his intended meaning but has made the proposition more explicit, more accessible to the reader.

Text 2.

Original: I personally believe that schools should be blamed for student suicides.

Revised: I personally believe that schools should be blamed for student suicides because as a student I feel pressure all the time.

If we look at the next example in Text 3 below, we can see that the writer has revised the original proposition by substituting 'cannot' with 'may not', thereby changing the meaning of the original proposition. Here, the proposition has been modified with the introduction of a more tentative tone by replacing the negative of the modal 'can' which denotes a high degree of certainty with the negative of the modal 'may' which denotes a lesser degree of certainty. Therefore, the writer has toned down or weakened the original proposition in the revised text. This is an example of propositional modification.

Text 3. (substitution: crossed out in the original text and given in bold in the revised text)

Original: Schools in foreign countries take you to an educational level that ~~cannot~~ be reached in Hongkong.

Revised: Schools in foreign countries take you to an educational level that **may not** be reached in Hongkong.

The next example is from a letter applying for a scholarship. Here again the writer has substituted 'very' with 'quite' and thus, by changing the degree of intensification, has modified the proposition.

Text 4. (substitution: crossed out in the original text and given in bold in the revised text)

Original: Moreover, I am ~~very~~ good at art and interested in it.

Revised: Moreover, I am **quite** good at art and am interested in it.

The next example, in Text 5 below, reveals a more complex propositional modification. The task was to write a letter to the editor expressing views about a new policy proposal. Here the writer begins by expressing agreement with the policy and then moves on to the problems she perceives with the policy together with possible solutions. One problem, according to the writer, is the price of the ticket. She has said so in the original. However, in the revised version the writer has modified the statement by using a modifier 'rather' and a prepositional phrase 'for students and poor families.' Here the writer has again modified the original proposition and the truth value of the revised proposition has changed. The truth value of the original version is that tickets are expensive for all categories of people. It is a generic statement with an absolute truth value. This means that it is applicable in all cases. The revised version, however, modifies the original proposition by weakening it. Now, instead of the proposition being applicable to everyone, it has been modified to apply only to the poor and to students. The truth values of the two propositions are **not** the same.

Text 5. (addition: in bold in the revised version)

Original: I think this is a very good method to keep the beaches clean. But...And also the ticket is expensive. We can reduce the amount to HK$30 not HK$50.

Revised: I think this is a good method to keep the beaches clean But...And also the ticket is **rather** expensive **for students and poor families**. We can reduce the amount to HK$30 not HK$50.

Writing on the same topic, we can see the writer of Text 6 has strengthened the original proposition by adding an adverbial to the sentence, the effect of which is to change the truth value of the original proposition.

Text 6. (addition: in bold in the revised version)

Original: I disagree with this proposal for the following reasons.

Revised: I **strongly** disagree with the proposal for the following reasons.

Examples from a graduate student

The following example is from a graduate student's thesis. The revisions were performed by the writer in response to her supervisor's comment.

In Text 7, the writer was advised by the supervisor to tone down the argumentation in order to match the genre of higher degrees academic theses where generalisations are presented in a non-assertive, cautious manner. This resulted in the writer revising her original propositions by weakening them considerably.

Text 7.

Original: As teachers of writing our aim is to help the writer find this power and the teaching of revision which will generate an awareness of audience, purpose and communicative intent is certainly one way of inculcating this voice.

Revised: As teachers of writing our aim is to help the writer find this power and the teaching of revision **as an attempt to** generate an awareness of audience, purpose and communicative intent **may be** one way of..

In the original text the truth values of the propositions are contained in two assertions. The first is that the teaching of revision will generate an awareness of audience, purpose and communicative intent. The second is that this teaching will certainly inculcate voice. In the revised version, the truth value of both of these propositions is modified. The first weakening occurs when 'will' is substituted by 'as an attempt to'. The revision is much more tentative and less assertive than the original. The truth value has been weakened; the **P** has been modified by a hedge. In the second revision, further hedging takes place when the strong assertion of 'will certainly' is much weakened by the substitution of the modal 'may be'. The Proposition **P** of the original is no longer same as the proposition **Q** of the revision. The strong assertion is not an assertion but a suggestion of a possible outcome. The function of the communicative intent has been changed. The truth values have changed.

Pedagogical and research implications

Propositional modification can be taught as a strategy for establishing and maintaining coherence in a text. As can be seen in Text 1, the student has modified her main proposition in the composition in order to match it to the existing content of the essay because in the body of the original composition a number of possible reasons for suicide were cited while in the main proposition only the schools were blamed (one cause only). The student has thus modified the original proposition in order to establish textual unity. She did not want to delete all the other causes of student suicide because they were valid reasons. She had to change the proposition, therefore, in order to make the text consistent throughout.

Problems of both understanding and teaching concepts of coherence are experienced by teachers and textbook writers. Some writers have tried to address the problem. Lauer et al. (1991:44-46) suggest that coherence is established by using the appropriate connectives, by maintaining the same reader role and by grammatical consistency. Lauer et al's suggestions, we feel, can be supplemented, along with other strategies, by teaching students how an original proposition can be modified in order to maintain consistency in a text.

Students are often taught ways of conveying moods, attitudes and feelings (White and Arndt 1991:157) by finding words which have the requisite association and implications, by using modal verbs indicating a sense of doubt, uncertainty, and levels of commitment by the use of expressions such as, **unfortunately**, and **luckily**, which signal writer attitudes. These writers are, in effect, suggesting that student writers modify their propositions. However, doing so in the creation of single texts is difficult. It is suggested that the opportunity to revise texts makes the application of these suggestions easier to implement.

This phenomenon of propositional modification also has implications for research into writing and revision. It is currently being used as a tool for revision analysis in research which investigates coherence in student writing (Sengupta, ongoing).

A final application of this phenomenon could be to increase the awareness of teachers of how discourse works, how coherence is created and how students can be helped to develop the cognitive and writing skills necessary for the development of meaning in writing.

Conclusion

The writers intend to continue exploring the nature and use of this phenomenon in student and adult writing in order to determine how widely this cognitive strategy is used and whether the findings of further research can be applied in the pedagogy of writing and in the area of text analysis.

Note.

The authors wish to thank Dr. Desmond Allison for comments on an earlier oral presentation of this paper.

References

Albrechtsen, D., Evensen, L. S., Lindeberg, A. C., & Linnarud, M. (1991). Analysing developing discourse structure: The Nordwrite project. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith & M. Swain (eds.), *Foreign/second language pedagogy research.* Clevedon, Philadelphia: Multilingual Matters.

Bell, R.T. (1991). *Translation and translating*. Harlow: Longman.

Callow, K. & Callow, J.C. (1992). Text as purposive communication: A meaning-based analysis. In W.C. Mann & S.A. Thompson (eds.), *Discourse description diverse linguistic analyses of a fund-raising text* (pp. 5-37). Amsterdam: John Benjamin Publishing Company.

Channell, J. (1993). *Vague language: A pragmatic analysis*. Oxford: Oxford University Press.

Copi, I.M. & Cohen, C. (1990). *Introduction to logic* (8th edition). New York: Macmillan.

Falvey, P. (1993). Towards a description of corporate text revision. Unpublished doctoral thesis, University of Birmingham.

Fredericksen, C.H., Lauer, J.M., Montague, G., Lunsford, A. & Emig, J. (1991). *Four worlds of writing*. Harper Collins Publishers.

Freedman, S.W. (ed.). (1986). *The acquisition of written language: Response and revision.* Norwood, NJ: Ablex.

Hurford, J.R. & Heasley, B. (1983). *Semantics: A course book*. Cambridge: Cambridge University Press.

Lauer, J.M., Montague G., Lunsford A. & Emig J. (1991). *Four worlds of writing*. Harper Collins Publishers.

Matsuhashi, A. & Gordon, E. (1986). Addition and the power of the unseen text. In S.W. Freedman (ed.), *The acquisition of written Language (rev. ed.)*. Hong Kong: University of Hong Kong.

Mann, W.C. & Thomson, S.A. (1986). Relational propositions in discourse. *Discourse Processes*, 9(1):57-90.

Mann, W.C. & Thomson, S.A. (1987). Rhetorical structure theory: Towards a functional theory of text organization. *Text*, 8(3):243-281.

Phillipson, R., Kellerman, E., Selinker, L., Sharwood Smith, M & Swain, M. (eds.). (1991). *Foreign/second language pedagogy research*. Clevedon, Philadelphia: Multilingual Matters.

Sato, T. (1990). Revising strategies in Japanese students' writing in English as a foreign language. Unpublished Doctoral Thesis, Indiana University of Pennsylvania.

Sengupta S. (Forthcoming). How secondary school teachers teach writing in Hong Kong secondary schools. Hong Kong: University of Hong Kong.

White, R. & Arndt, V. (1992). *Process writing*. Harlow: Longman.

WRITING FROM SOURCES: DOES SOURCE MATERIAL HELP OR HINDER STUDENTS' PERFORMANCE?

Jo Lewkowicz

Introduction

There appear to be two basic reasons why the task of writing from sources is included in proficiency tests such as IELTS, TEEP and OTESL which are designed to assess students' ability to cope with language for academic study. First, it is argued that providing students with texts on which to base their writing ensures that they have something to say and that no student is disadvantaged through lack of information. As Weir (1993:135) has pointed out:

> By basing writing tasks on written and/or spoken text supplied to candidates or on non-verbal stimuli, it is possible to ensure that in terms of subject knowledge all start equally, at least in terms of the information available to them.

Secondly, and perhaps more importantly, such writing purportedly replicates the writing students are expected to undertake in their academic studies outside the language classroom. As Weir (1983:378) found in determining the tasks to be included in the TEEP test

> ... the task most students have to cope with across a range of disciplines would be the selective extraction of relevant information from a written corpus of information and verbal input and the subsequent reformulation of data from either or both sources in a piece of extended writing.

Thus, as Read (1990:113) has noted, providing test-takers with content material on which to base their writing

> may help reduce the effects of differences in background knowledge among test-takers and, when the writing tasks are linked with earlier reading and listening tasks, may represent a better simulation of the process of academic study than simply giving a stand-alone writing test.

But the effect of providing students with source material on which to base their writing has not been fully investigated. It is conceivable that such source text(s) impede rather than enhance performance and that students would write better without them. Their appropriate use depends on students' understanding the texts provided and this, in turn, may depend on their background knowledge of the subject under consideration. It is possible that students lacking sufficient depth of knowledge have to rely heavily on the background texts and that they may inadvertently be 'encouraged' to replicate not only the content but also the lexis and structure of the original texts.

Writing from sources, nevertheless, is an authentic task in that it is required of many tertiary students across a wide range of disciplines (Weir 1983, Horowitz 1986). It involves skills which students need to master if they are to be accepted into an academic community. It is therefore a valid test-task both in terms of face and content validity. Yet it is a task which many students, particularly those whose L1 is not English, appear to find difficult. It is a task that needs to be more accurately described and fully understood if it is to be valid not only in terms of face and content validity but also in terms of construct and predictive validity.

There has been considerable interest in the reading-writing relationship (see, for example, Esterhold 1990, and for a comprehensive review, Tierney & Shanahan 1989). There has also been much research into the writing process both in the L1 (see among others, Britton et al. 1975, Graves 1983), and the L2 (including studies by Zamel 1983, Raimes 1987 and Cummings 1989). But little is know about how successful writers integrate sources into their own writing or the difficulties that non-native speakers encounter in dealing with such tasks. Those studies which have dealt with the synthesis of information from sources have focused either on writing a summary from a single source (e.g. Taylor 1984) or on the effect of reading ability on writing (e.g. Kennedy 1985). They have failed to look at how the information from a text is used in the students' own writing.

One notable exception to this is the study by Campbell (1990) who compared the way in which native and non-native speaker students used a background reading text to write an academic essay. She found that although both groups have shown an ability to use the source text appropriately, that is to quote from it, paraphrase, summarise and explain it in their own writing, they still relied heavily on copying, not only in the appropriate form of quoting from the text, but also in the inappropriate forms of exact or near copies from it. She also found that both groups frequently failed to reference the text and used it more frequently to foreground information, that is, to make a point rather than to support their own content.

A serious limitation of this study was that the students' writing was based on a single background text. In academia, students are generally expected to assimilate information from a number of sources, some of which may present conflicting points of view. According to Campbell, using a single source may have contributed to the students' failure to attribute since they were aware of the fact that their instructors were familiar with the text.

> If these same students were given another writing task involving the use of a number of sources presenting conflicting views, they might have provided more documentation to clarify the sources of the various views. (Campbell 1990:223)

Furthermore, although it is common to assign essays based on background reading, in contrast to the instructions given by Campbell, instructions to students do not usually specify the source(s) on which the answer is to be based. Providing such specific instructions may have contributed not only to the lack of attribution, but also to the extent and location of the text references in the students' essays.

Aim of the Study

The aim of this study was to compare, given the same task, the writing of non-native speaker students provided with background reading texts for their own academic writing with that of students who had not been given the texts. The quality of the essays was compared in terms of length, the number and extent of elaboration of the points presented, as well as the overall success (or failure) of the writers to substantiate a point of view. In addition, for the group who were given the texts, the way in which the information was used and the extent of attribution to authors of the background texts were also recorded.

Subjects

The 75 non-native speaker subjects for this study were enroled in the English for Arts Students (EAS) course at the University of Hong Kong. This is a compulsory EAP course for all first year students in the Faculty of Arts at the University. The students were from six of the forty groups being taught, the students being randomly selected into groups at the beginning of the academic year. In terms of *Use of English*, which is a public examination held at the end of S7 (Grade 13), all students entering the Faculty have attained a minimum grade 'D8' (maximum 'A1'), with a small number of mature students being granted an exemption from the examination.

Materials and Method

Three of the groups participating in the study were selected to do the task with the texts (Group 1) and three without the texts (Group 2). The mean scores for the students' *Use of English* examination were compared across the two groups to confirm that there was no difference in language proficiency between those being assigned the task with the texts and those being assigned the task without the text (Table 1).

For the three subgroups with the texts (Group 1), instructors were asked to give these out to students 25 minutes before the end of a class. The students were informed that they would be using the texts on a subsequent occasion, though they were not given details of the writing task at this stage. The students had time to read the texts, make notes on the side and generally think about the subject. The texts the students had worked on were then collected in and the same texts with the student's notes were returned to individual students at the beginning of the following class (no later than 2 days after the first class). These students were then given the same essay question as those in the three subgroups (Group 2) who were selected to complete the task without the texts. There were in total 32 students working with the texts and 43 without the texts. All the students were given a full class (50 minutes) to write the essay.

Table 1

Differences in mean UE scores for the two groups

	mean	s.d.
Group 1 (n = 29)	6.276	1.386
Group 2 (n = 40)	6.125	1.771
		t=.38, p=.704 (n.s.)

Note: n = 69 : UE for 6 students not available

The students working with the texts were given two extracts on Hong Kong identity, one from an article (Baker 1983) and another from a book (Yee 1989). The extracts, which were of approximately the same length, i.e. the former of 745 words and the latter of 805 words, were selected because of their relevance to the essay topic and because they had been piloted with previous groups of students at the University. They also presented similar information but from a different perspective, not always coming to the same conclusions.

With 1997 approaching and the impending return of Hong Kong to China, students in Hong Kong are concerned about the issue of preserving their own identity. Hence it was believed that they would be able to write something within the allocated time on the subject even without the texts. However, it was hoped that the texts would complement what the students had to say and would enable them to support their ideas. The full essay topic was as follows:

> To what extent do you believe that Hong Kong people have a unique culture and identity of their own? Give evidence to support your views. You may wish to consider some or all of the following points.
>
> - Historical ties with China
> - Family life

- Life-style
- Interests/pastimes
- Work styles and habits

The prompts, which cover the main areas discussed in the texts, were given so as not to disadvantage the students without the texts. There is also evidence that such questions are set outside the language classroom and not only at the University. (See, for example, Horowitz 1986, for tasks set by subject professors in the U.S.)

The essays were marked holistically by a class instructor on a scale of 1 to 10. The researcher then double-marked all the student scripts. (In a follow-up study, the same scripts will be analytically marked to determine whether any specific differences in writing quality across the two groups can be discerned.)

Once the essays had been typed out for ease of reference, they were analyzed in terms of the number of points introduced by each of the writers. First the points were counted for the two texts, text 1 introducing 22 different points and text 2 introducing 11. (These points can in effect be regarded as linguistic propositions in that they are assertions that contain truth value, though many of them in text 1, for instance, were listed rather than developed, explained or elaborated.) For Group 1 students each of the points put forward were matched against those identified in the texts. Then any additional points introduced by the students were recorded. For Group 2, the total number of points supporting the students' arguments was noted.

Finally, one of the subgroups of students (11 students in total) who did the task with the texts was interviewed to better understand how they perceived the task and why they used the texts as they did.

Analysis of Data

The mean scores awarded by the instructors and the researcher for the two groups of students were compared to see whether either of the groups performed significantly better than the other. In addition, the mean lengths of the essays in terms of the number of words were compared across the two groups as it was hypothesized that students with the texts would have more to say about the culture and identity of Hong Kong people than those without the texts. Finally, the mean number of points introduced by each group were compared to verify whether the background texts did provide Group 1 students with additional information to use in their writing.

For the group with the texts, the mean ratio of points generated from each of the texts was calculated to see whether both texts had been used equally. The mean ratio of points from the texts to student-generated points was also estimated. Then the points which students had used from the sources were examined to see whether they had been used as background or foreground information, and whether they had been acceptably attributed to the relevant author.

Results and Discussion

The interrater reliability of .61, though significant at $p < .001$, was not high. However, the difference in mean scores for the essays across the two groups of students showed no significant difference at the .05 level (Table 2).

Table 2

Differences in mean task scores for the instructors and the researcher for the two groups

	Instructors		**Researcher**	
	mean (max 10)	**s.d.**	**x (max 10)**	**s.d.**
Group 1 (background texts	6.219	1.845	6.219	1.641
+ essay) (n = 32)	6.395	1.904	6.302	1.897
Group 2 (essay only) (n = 43)				
	t = .40, p = .688 (n.s.)		t = .20, p = .842 (n.s.)	

Thus, though there was considerable variability in how the raters assessed individual essays, neither the instructors nor the researcher appear to have perceived a difference in standard of writing between the two groups. The group without the texts seems to have been equally as successful in fulfilling the task as the group with the texts.

This finding is further supported by the fact that though there was some difference between the mean length of the essays across the two groups, this difference was not significant at the .05 level (Table 3).

Table 3

Differences in mean length of essays for the two groups

	No. of words	
	mean	**s.d.**
Group 1 (background texts + essay) (n = 32)	435.188	122.146
Group 2 (essay only) (n = 43)	397.372	100.841

$t = 1.47$,
$p = .147$ (n.s.)

Since one of the reasons for providing students with texts is to ensure they have something to say, it was assumed that Group 1's essays would have been more detailed. This, however, appears not to have been the case. Not only was there no significant difference between the mean lengths of the essays, but the range for the two groups was also very similar. Group 1 students wrote between 221-731 words while Group 2 between 196-694 words. It must, however, be remembered that the amount students wrote may have been limited by the time restriction. Students in Group 1 may have written more if they had been given more time, a fact confirmed by some of the interviewed students. In the same way, students in Group 2 may have written more if they had had more time to think about the subject: they were not given the same opportunity as Group 1 to consider the subject in advance.

The one difference between the two groups that was significant at $p < .05$ was the mean number of points students introduced to support their argument (Table 4).

This suggests that the texts helped students to generate additional ideas, the ratio of ideas found in the texts to student generated ideas not found in the texts being approximately 4:1. But since there was no significant difference in the mean length of the essays, Group 2 students without the texts appear to have developed their ideas more fully. In itself this can be seen as a positive rather than negative trait, a point made by Falvey (forthcoming) who, in discussing the attributes of a good essay, noted:

> In an essay, the force of a proposition and the effect of it on a reader is strengthened if the proposition contains supporting arguments, or if it is elaborated or explained further.

Table 4

Differences in mean number of points introduced for the two groups

	mean	s.d.
Group 1 (background texts + essay) (n = 32)	10.625	3.722
Group 2 (essay only) (n = 43)	8.605	3.375

t = 2.45,
p = .017*

*significant at p<.05 level

Although there was some overlap in the ideas presented in the two texts, the ratio of ideas between text 1 and text 2 was 2:1. If the students in Group 1 had used the texts equally, one would have expected the same ratio of points to be taken from each text. In reality that proved not to be the case; the ratio of points used from text 1 and text 2 was approximately 9:2. This would imply that text 1 was easier for the students, but using the Fog index (Gunning 1952) it was found to be the more difficult passage. With a Fog index of 17, text 1 could, in fact, be considered too difficult for undergraduate students, while text 2 with a Fog index of 15 at about the right level. An alternative explanation, and one confirmed by the students when interviewed, was that text 1 lent itself more readily to the direct lifting of ideas from the text. According to some of the students, Baker 1983 (text 1) was easier to copy while Yee 1989 (text 2) was more difficult as the relevant ideas in the latter were scattered and the text included a number of points that were not related to the question.

The extent of exact and near copying among the students in Group 1 appears on first analysis to be extensive (though more detailed analysis remains to be undertaken in a subsequent study). All 32 students used ideas from one or both passages, yet only 6 made any attempt to attribute and of these only 1 attributed and gave the references at the end of the essay. Campbell (1990) in her study justified the lack of attribution by the fact that her students were given only one text and they knew their instructors were familiar with it. In this study, the fact that the students were given two texts appears to have made little difference. Even though the students had been taught to attribute in their EAP classes, few demonstrated an ability to do so, and even fewer an ability to do so correctly. Examples of attribution from the students' writing include the following:

> As in Yee, A.H.'s passage, the pursuit of material interests could be a show of family ties ... (Student W101/1/10)
>
> As Mr Baker said in his book, the emergence of Hong Kong man, China Quarterly, the common theme of Hong Kong life is "Drink your wine today and be drunk". That means Hong Kong people are ... (Student W108/5/23)
>
> For most of these people, Hong Kong is only a temporary place of residence and work. (Baker 1983). "It was a business location to be exploited and then to be retired from in later life, ...". (Student W101/9/18)

In the few essays where the attribution was correct it was not always consistent. For example, in essay (W101/4/13) ideas taken from text 1 made up approximately one-third of the essay, yet the writer only made reference, albeit correctly, to Baker twice.

Campbell (ibid.:221) in her study also suggested that there was no evidence that students intentionally or through lack of knowledge violated the accepted conventions of attribution. According to her they may have failed to reference simply because of time pressures.

> Given time constraints and the physical constraints of a full classroom, most writers find it difficult to produce quality writing. To include information from written sources without violating conventions of acceptability is even more difficult. The possibility exists that some of these students might have even intended to eliminate Near Copies in later revisions of their papers.

In the writing of Hong Kong students there is little evidence that they would 'eliminate Near Copies' (i.e. parts of their writing which closely adhered to the wording of source texts.) Teachers of EAP and subject-specialists have long complained that students are not averse to plagiarism and the interviewed students themselves said it was easier to copy as in that way they could avoid mistakes. They did not see anything wrong in this as they claimed they were copying from authorities who would inevitably be better than themselves. This is supported by a study by Deckert (1993), who found that Hong Kong Chinese tertiary students, brought up in a tradition of rote memorisation and an adherence to a limited number of 'authorities', appear to have a very different attitude to plagiarism than their Western counterparts. Whereas

> American freshman registered concern for authors, fellow students, ownership rights, and responsibility for doing independent work, the first

> year Hong Kong students seemed to approach only the latter concern and then only with a pronounced egocentric emphasis. That is, they viewed plagiarism to be wrong because it hampers their own learning and disturbs their sense of personal integrity. The other issues seemed to be of little consequence. (Deckert 1993:141-142)

A careful examination of the way in which students used the ideas from the texts further demonstrates their tendency to produce these ideas as their own. It also indicates, as can be seen from the examples below, their inclination to select key words from the texts and, often out of context, string them together without any development or explanation, making their points at times difficult to follow. (The italics in the examples indicates the parts taken from the texts; the errors are those made by the students.)

Example 1

> Besides, *the major religion of settled China has been ancestor worship, which is associated with continuity of family and land ownership, but the religions of Hong Kong Chinese are those which stress movement and expediency.* So you can find that *there are a large number of Christians in Hong Kong,* but on the other hand *many people go to temples whose gods have the reputation of bringing fortune and happiness, and when the prayers are not answered they move on to another god or temple.*
>
> Many foreigners said the work styles of Hong Kong people are distinctive, they are *highly competitive, tough for survival, quick thinking and flexible.*
>
> Hong Kong people like to wear *western clothes, speaks English or expects his children to do so, drinks western alcohol* ... etc. They are not westerners, but *almost alone in the Chinese world Hong Kong has not adopted Mandarin as the lingua franca,* it is really unique. (Student W108/14/32)
>
> [Note the order of points follows that in text 1, but the writer has been selective in the points chosen. Integration of the text with the student's own writing fails in the latter part. The section quoted makes up approximately one-quarter of the essay.]

Example 2

> In addition, HK people prefer western life style to Chinese one. HK people like wearing *western clothes, speak English and expect exciting* as well as new life-style.
>
> Furthermore, HK people develop a unique work styles and habits. *Quick thinking, flexible* and *competitive* attitude is the main working feature. *Fast*

working *pace* was adopted in life-style. Also, Hong Kong people pursue immediate return and satisfactions making people seem to be more *materialistic*. (Student W136/4/4)

[Note that the writer in this short section uses a variety of key words from the second and third paragraphs of text 1 and makes minor changes such as 'quick satisfaction' to 'immediate return and satisfactions'. In addition, the writer's failure to develop ideas, for example in the first paragraph, makes some of the points difficult to follow. What does he mean by '... speak English and expect exciting as well as new life-style'?]

Example 3

Deep westernization is another aspect of Hong Kong culture. Hong Kong people are *go-getting, highly competitive, tough for survival, wear western clothes, speak English, drink alcohol, have sophisticated tastes in cars, and household gadgetry. But they are not westerner, they are merely westernized.* Meanwhile *they are not Chinese in the sense that the citizens of the People's Republic of China are Chinese.* They are not westerner but westernized, not the citizens of China but Chinese. It is their identity. (Student W101/5/14)

[Note how the writer uses the plural rather than the singular of the original text, but otherwise copies from text 1. This section makes up approximately one-third of the total essay.]

A further characteristic of the students' writing also emerges from the essays. Some of the students who were able to summarise the main points of the texts were also able to use such information as background in support of what they were saying, as in the example below:

Example 4

Besides, the working attitude of Hong Kong people, their pastimes and interests are also similar. Today, Hong Kong's economy has been greatly improved and hence, Hong Kong people can and are willing to spend money in luxuries and high-class services. It is doubtful whether any other country or city will have so many luxury advertisements broadcast every day, from diamond watches to first-class cars. (Student W136/3/3)

[Note that the student's main proposition is Hong Kong peoples' interest in money and how she uses the idea of Hong Kong people having 'sophisticated tastes' as an example of that interest. She is using information from the text to support her own ideas, i.e. as background rather than as foreground information.]

However, in line with Campbell's (1990) findings, the majority of students appear to have used the texts as sources of foreground information, to introduce new propositions rather than to support their own ideas (see examples 1-3 above). The reason for such foregrounding may lie partially in the texts themselves. Both were short and packed with information that was not extensively developed. As one of the interviewed students pointed out, if the points needed explanation, the authors of the texts would have given the necessary detail.

Conclusion

The purpose of this study was to determine whether giving students texts upon which to base their writing enhances their performance and if so, in what way(s).

The initial results reported here indicate that background texts provide student writers with ideas, but they do not necessarily improve the quality of their writing. Given texts, students appear to rely heavily on these and less so on their own ideas. They do not develop their propositions fully which may make their writing difficult to follow. In contrast, students' writing without access to background texts support their ideas more extensively, even though they may have fewer ideas to contribute.

What seems to distinguish the more successful writers among those with the texts is their ability to summarise the main ideas in the passages and then integrate them in their own writing. This seems to be closely related to the skill of backgrounding information-a skill which many first-year undergraduates in Hong Kong find difficult. It may be that this skill needs to be explicitly taught (and tested) if students are to be weaned away from foregrounding source-text information in their own writing.

Dependence on background texts appears to lead to the breaking of academic conventions as well as the foregrounding of text information. It may hamper students' writing rather than enhance it. Thus, the argument that students' writing without text support are disadvantaged remains questionable. The main justification for maintaining text-based writing tasks within EAP courses and tests remains that such tasks replicate those students have to cope with in their studies: they are a stepping-stone to being able to cope with academic writing. Yet, since students probably have to deal with various forms of non-text based writing outside the university setting, it may be that such non-text based writing should also be encouraged and tested.

References

Baker, H. (1983). Life in the cities: The emergence of Hong Kong man. *The China Quarterly*, 95:469-479.

Britton, J., Burgess, T., Martin, N., McLeod, A. & Rosen, H. (1975). *The development of Writing Abilities* (pp. 11-18). London: Macmillan.

Campbell, C. (1990). Writing with others' words: Using background reading text in academic composition. In B. Kroll (ed.), *Second language writing: Research insights for the classroom*. Cambridge: Cambridge University Press.

Cummings, A. (1989). Writing expertise and second language proficiency. *Language Learning*, 39(1):81-141.

Deckert, G. D. (1993). Perspectives on plagiarism from ESL students in Hong Kong. *Journal of Second Language Writing*, 2(2):131-148.

Eisterhold, J.C. (1990). Reading-writing connections: Toward a description for second language learners. In B. Kroll (ed.), *Second language writing: Research insights for the classroom*. Cambridge: Cambridge University Press.

Falvey, P. (forthcoming). Section C: Writing. In E. Marshall, J. Fullilove & C. Lee (eds.), *Teaching for the use of English examination*. Hong Kong: Hong Kong Examination Authority.

Graves, D. (1983). *Writing: Children and teachers at Work*. London: Heinemann Educational Books.

Gunning, R. (1952). *The techniques of clear writing*. New York: McGraw-Hill.

Horowitz, D. M. (1986). What professors actually require: Academic tasks for the ESL classroom. *TESOL Quarterly*, 20(3):445-462.

Kennedy, M. L. (1985). The composing process of college students' writing from sources. *Written Communication*, 2(4):434-456.

Raimes, A. (1987). Out of the woods: Emerging traditions in the teaching of writing. *TESOL Quarterly*, 25(3):407-430.

Read, J. (1990). Providing relevant content in an EAP writing test. *English for Specific Purposes*, 9:109-121.

Taylor, K. K. (1984). The different summary skills of inexperienced and professional writers. *Journal of Reading*, 27:691-699.

Tierney, R. J. & Shanahan, T. (1991). Research on the reading-writing relationship: Interactions, transactions and outcomes. In R. Barr, M. Kamil, P. Mosenthal & P.D. Pearson, *Handbook of Reading Research Vol. 2*, pp. 246-280. New York: Longman.

Weir, C. (1983). Identifying the language needs of overseas students in tertiary education in the United Kingdom. Unpublished PhD Thesis, University of London Institute of Education.

Weir, C. (1993). *Understanding and developing language tests*. London: Prentice Hall.

Yee, A. (1989). *A People misruled: Hong Kong and the stepping-stone syndrome.* Hong Kong: Chinese University Press.

Zamel, V. (1983). The composing process of advanced ESL students. Six case studies. *TESOL Quarterly*, 17:165-187.

EFFECTS OF SIGNALLING ON READING COMPREHENSION

Icy K.B. Lee

Introduction

In the area of text processing there are a number of studies which have focused on a group of metadiscoursal words and phrases that are called "signals". Signals are structural cues that specify the interrelationships among items of information which compose the text and indicate the subordination and coordination of this information (Meyer 1984a:117). They are thought to indicate to readers the importance of ideas in a text, thus facilitating comprehension. The present study investigates the effects of headings, previews, and logical connectives on secondary ESL (English as a second language) students' comprehension of expository prose. Headings are short phrases that announce superordinate content (main ideas) before the reader encounters the actual content. Previews are phrases or sentences that announce superordinate content and relationships among superordinate content before the reader encounters them. Logical connectives are words or phrases that interrelate superordinate or subordinate content (details) to content of similar or different hierarchical levels (Spyridakis and Standal 1986). Since the majority of Hong Kong secondary students are not skilled readers, if consistent effects for signals are established, methods can be found to help these less skilled readers with their reading. Moreover, texts can be written with thoughtful inclusion of signals to facilitate ESL readers' reading comprehension, not just in the English language subject, but also in other content subjects taught in English.

Various signal types have been identified in reading research, such as headings, titles, previews, topic sentences, conjunctions, logical connectives and pointer words. Some studies have examined the effects of single signal types (e.g. Hartley and Trueman 1983, Klare, Shuford and Nichols 1958) while some have examined multiple signals in combination (e.g. Loman and Mayer 1983, Meyer 1975, Walker and Meyer 1980). Insignificant effects were found for structural cues, previews, summary statements and pointer words on recall of expository prose (Britton, Glynn, Meyer and Penland 1982, Meyer 1975, Meyer, Brandt and Bluth 1980). The inadequacy of recall, pointed out by Loman and Mayer (1983), lies in the fact that it only measures the overall amount recalled, hence a quantitative approach. In their study, Loman and Mayer use a qualitative approach which measures the quality of recall. Results show that headings, previews, and logical connectives help both good and poor readers recall conceptual ideas and aid creative problem solving, i.e. inferencing. It is concluded that signals would have greatest effects for unfamiliar technical or scientific prose (Loman and Mayer 1983:410). Kintsch and Yarbrough (1982) also adopt a qualitative approach to measure the effects of signals, using cloze and direct questions, and significant effects were found for signals in combination. Previews were shown to help good readers especially in comprehending longer, difficult and unfamiliar texts (Spyridakis and Standal 1986). Evidence of the significant effects of headings, previews, and logical connectives on comprehension of details and inferencing is available in Spyridakis and Standal (1987 and 1989b), but such facilitation seemed to depend on text length and text difficulty. That is, signals would have clearest effects when the texts were neither too easy nor too difficult (Spyridakis and Standal 1987). Previews were found to be more helpful than logical connectives overall (Syprdiakis and Standal 1989b). Also each signal type individually was found to aid readers' comprehension more than any two in combination, e.g. headings are most helpful when they are alone (Spyridakis and Standal 1989b).

Although we still cannot find conclusive evidence for the effects of signals, certain important points have emerged from previous studies. Previous research on signals has indicated that if a text is syntactically or lexically quite difficult, the reader may benefit by the presence of signals. If a text is easy, on the other hand, readers will not benefit from signals because they simply do not need help from them. Neither would readers benefit from signals if the text is extremely difficult because they "may cause overloading of the subject's processing system and cause complete or near-complete breakdown of the comprehension process" (Afflerbach and Johnston 1984:314). Furthermore, signals seem to aid readers best when the topic of the text is unfamiliar (Loman and Mayer 1983, Spyridakis and Standal 1986, 1987, Spyridakis 1989a, 1989b).

The present study continues the previous line of research and examines both the individual and combined effects of headings, previews and logical connectives, using a relatively difficult passage on an unfamiliar topic. The research question addressed is: would signals benefit less skilled ESL readers in comprehending relatively difficult expository texts at the superordinate level (i.e. main ideas), and how? Since signals have been found to help readers identify superordinate content (Spyridakis 1989b), the present study mainly concentrates on comprehension of superordinate content and inferencing made from it. The construct "comprehension" is operationally defined in terms of the comprehension test scores. There are two dependent variables: (1) factual comprehension and (2) inferential comprehension. The independent variables are the three signal types that are investigated in the study.

The study, however, differs from the previous studies in a number of ways. First of all, the majority of these signalling studies were conducted with first language students, but this study investigates the effects of signals on some ESL learners' reading comprehension. Secondly, while most signalling studies have used good comprehenders studying in colleges, the subjects in this study are secondary students identified as less skilled readers. Thirdly, technical prose, which has often been used in previous studies, is not used in the present study because it will be too difficult for Hong Kong secondary students and will probably be at their frustrational level of reading. Instead, a non-technical expository text is used. Finally, instead of a recall test, multiple-choice test or forced-choice test, a comprehension test requiring written answers from students is used, which, it might be argued, is a more direct measure of comprehension. It is also a qualitative approach which allows one to examine and analyze the answers in greater depth.

Method

Subjects

Altogether 285 female S4 (secondary four, i.e. Grade 10) students, aged between 15-17, took part in the study. They are from three well-established girls' schools in Hong Kong with Band 1 to Band 3 intake at S1 (Grade 7) (altogether there are 5 bands, Band 1 being the best). The "less skilled readers" were selected from the weakest classes in the three schools, where students were expected to get Grade D or E in English Language (Syllabus B) in the Hong Kong Certificate of Education Examination in S5 (Grade 11).

Materials

An expository passage on an unfamiliar topic was chosen for the study. The passage belongs to the discourse type "collection of descriptions" in Meyer's (1984c) classification, which refers to a text where "a number of attributes, specifics, or settings are given about a topic" (Meyer 1984b:123). It contains information about the various uses of the aeroplane, and it is the kind of passage students may encounter. However, the topic, namely general aviation, is quite unfamiliar to secondary students, and it is not covered in their content subjects such as Geography and Economics. In determining the difficulty level of this passage, expert opinions were sought from several experienced secondary English teachers, who read the passage and expressed similar views that the passage would be quite difficult, but not too difficult, for the ability group identified in the study. The passage is about 500 words, which would be considered a rather long passage by S4 (Grade 10) ESL students.

Passage Analysis

The superordinate content of the passage was identified by having two independent judges, an expatriate teacher trainer and a local school inspector listing the main ideas of the passage. The judges' analyses were similar. Since an agreed list of main ideas was needed for operational purposes in the research, minor differences were resolved in conference.

Passage Design

The non-signalled control passage was derived from the authentic passage on "General Aviation" (see Appendix 1) with all existing signals belonging to the categories of headings, previews and logical connectives deleted. Then, headings, previews and logical connectives were added to the passage, either singly or in combination to produce seven other versions of the passage. Altogether six headings and six previews were added to the signalled passages to pre-announce the superordinate content. Twenty logical connectives, derived from Halliday and Hasan's (1976) list of conjunctive relations (e.g. *In fact, on the other hand, as a result, however,* etc.) were added to the signalled passages, placed among ideas at all levels to interrelate superordinate and, subordinate ideas. The eight versions of the passages either had:

1 no signals;
2 headings only;
3 previews only;
4 logical connectives only;
5 headings and previews;
6 headings and logical connectives;
7 previews and logical connectives; or
8 headings, previews, and logical connectives.

The length of the passages ranged from 429 to 519 words. (See Appendix 2 for the signalled passage.) Three different readability tests, Flesch Reading Ease, Gunning's Fox Index and Flesch-Kincaid Grade level, were used to ensure that the eight versions of the passage were of comparable difficulty level. No significant differences were found. (Full details are reported in Lee 1993.)

Test Design

The test comprised a factual question asking about the five main ideas in the passage and five inferencing questions (see Appendix 3). These questions were designed to assess two types of information:

1 comprehension of superordinate content, and
2 inferences made from superordinate content in the passage.

Two other questions were asked before the comprehension test to obtain information about subjects' familiarity with the topic and their view of the difficulty level of the passage.

Procedure

The passage was pilot-tested twice with two different groups of students, where they were asked to rate the familiarity and difficulty level of the passage. The passage was found to be unfamiliar and relatively difficult. The test was then administered to the 285 subjects in a regular English lesson. Each subject was randomly assigned to read one of the eight versions of the passage. Subjects were given a set of materials, which included a cover page with simple instructions, one passage, a question on topic familiarity, a question on passage difficulty and a comprehension test on the passage. Verbal instructions were given by the teacher before the test. Subjects were given 10 minutes to read the passage and answer the familiarity question and the difficulty question first. They were then given 25 minutes to answer the comprehension questions. They were allowed to refer back to the passage.

The tests were scored by two raters. Pearson Correlation Coefficients were calculated between the scores assigned by the two raters for each dependent measure. Very strong positive correlations were found for both the factual (0.9863) and inferential (0.97886) comprehension scores.

The test data were analyzed with SAS (Statistical Analysis System) statistical software for each dependent measure (factual comprehension and inferential comprehension scores), using the GLM (General Linear Models) Procedure. One-way ANOVAs were used to find out the single and combined effects of signals. Three-way ANOVA was used to investigate the main effects of signals and their interactions. A significance level of $p = .05$ was chosen.

Results

285 subjects participated in the study, but only 232 scores for each dependent measure were submitted for analysis. The scores of 53 subjects were removed from the data base, they were those who (a) indicated the content was not unfamiliar to them (39 subjects), (b) did not answer the familiarity question and/or difficulty question (9 subjects), and (c) indicated that the passage was very easy or easy (5 subjects).

The mean scores and standard deviations for the eight versions of the passage were calculated (see Table 1 in Appendix 4). The descriptive statistics indicate a general trend toward higher scores with the addition of signals for factual comprehension, particularly in passages where headings were added. However, the inferential comprehension scores for all the eight versions seemed to fall within a small range. If we examine the standard deviations, we can find that the dispersion of the factual comprehension scores was greater than that of the inferential comprehension scores.

Results of One-way ANOVA on individual signals showed insignificant effects for inferential comprehension. But there were significant results for headings ($p = .0039$) and previews ($p = 0.0120$) for factual comprehension (see Table 2 in Appendix 4).

Insignificant effects were also found for inferential comprehension for all the three versions of the passage containing two signal types. There were significant effects, however, for all the three versions of the passage for factual comprehension with headings and previews ($p = .0001$), headings and logical connectives ($p =$

.0111), and previews and logical connectives (p = .0444). In other words, there were significant combined effects for these signals (see Table 3 in Appendix 4).

Results of one-way ANOVA on the 3-crossed effect of headings, previews and logical connectives indicated insignificant results for inferencing, but there was significant effect (p = .0001) for factual comprehension. When one examines the results of the previous analyses, one can notice the highly significant effect of headings on factual comprehension. It may therefore be possible that the significant three-crossed effect (headings x previews x logical connectives) was due to the very predominant influence of headings (see Table 4 in Appendix 4).

Results of the 3-way ANOVA (headings x previews x logical connectives) (see Table 5 in Appendix 4) indicated main effects of headings and previews for factual comprehension. There was no interaction, however, between headings and previews, and there was no other interaction for the other signals.

Discussion

It is hypothesized that signals would aid comprehension at the superordinate level not only when they are alone but also when they are combined. The individual and combined effects of signals were tested in the study and analyzed in the factorial design. Two significant main effects were identified for individual signals in the statistical analyses. Headings and previews were found to improve factual comprehension at the level of superordinate ideas. There was, however, insignificant effect for logical connectives for both factual and inferential comprehension. No significant results were found for inferencing for any of the signal types. Of importance is that there was no significant signal interaction in the analyses. The fact that there was no interaction between headings and previews can be explained by the similarity of their functions, both of which serve to pre-announce superordinate content.

The significant results of headings and previews for factual comprehension are consistent with previous research findings. Since headings pre-announce the superordinate content before the readers encounter it, they help readers locate the superordinate ideas and relationships more easily, facilitating the comprehension process. The influence of previews on factual comprehension could be explained by the selective attention and memory activation hypotheses put forward by Glover, Dinnel, Halpain, McKee, Corkill and Wise (1988). That is, previews selectively guide readers' attention and help their memory activation. Like headings, previews can help readers build a strong hierarchical framework in memory from which comprehension of superordinate content is enhanced.

However, the significant effects of headings and previews on factual comprehension demonstrated in this study may have been due to other reasons. First of all, headings and previews might be particularly helpful to students reading this text because of its "list-like" nature (i.e. "collection of descriptions" type). Since the passage was a collection of ideas about the uses of general aviation, with the headings and previews virtually summarising the main points, students were able to get the superordinate ideas from the headings and previews quite easily. The greater influence of headings than previews on factual comprehension might have been due to the specific question in the test. It could be possible that the question which asked subjects to list *the* five main uses of general aviation in the passage had led students (who were reading a passage with headings) to look for the clues in the headings. Also headings might have helped comprehension better than previews because they attracted attention more easily, being short phrases added to the passage and being underlined, and hence more striking visually. In brief, headings might have been the best aid to comprehension because of the clouding effects of the text type and question type.

The insignificant effect of logical connectives in the study is consistent with the findings in Spyridakis and Standal (1986), who have attributed the result to the fact that logical connectives did not support or summon the readers' structure strategy when they were reading a relatively difficult passage. This explanation is highly plausible in the present study, where students were also reading a relatively difficult passage. Another possible explanation for the result is that since logical connectives were added at many levels of the passage (unlike headings and previews, which were added at the superordinate level only), readers might have been distracted by those logical connectives that emphasized subordinate information and relationships when they were building a hierarchical representation of the text in memory (Spyridakis 1989b, 408).

Apart from these reasons, it may be possible that logical connectives in the passages were simply not heeded by the subjects. Allison has pointed out that explicit signals of relations in a discourse are often not heeded by text receivers (Allison 1991:378). A number of studies have reported that ESL readers have difficulties handling cohesive links (Cohen et al. 1979; Dubin and Olshtain 1980). The insignificant results of logical connectives may have been due to the fact that students in the study simply ignored the logical connectives when they encountered them, or they misinterpreted them in the process of comprehension. Crewe (1990) has cited an example where a reader was reading "therefore" in the text and yet expecting "however". In such a case, the reader would simply overrule the original word and read "however". One could perhaps conclude that logical connectives did not aid comprehension because readers did not pay adequate attention to them, or they misconstrued some of the logical connectives in the comprehension process.

Related to the above is the observation that logical connectives did not help students of this particular standard. Geva's (1992) study on conjunctions with L2 learners shows that L2 learners improve their ability to utilize and infer logical relationships in extended discourse as they grow in their L2 proficiency. Grabe also points out that there is a language proficiency threshold below which comprehension processes used in L1 reading are not used as effectively in L2 reading, and therefore language proficiency plays a crucial role in second language reading abilities (Grabe 1991:391). Given the students' standard of English, the logical connectives used in the study, which were of a fairly wide range (19 different logical connectives altogether), would probably be beyond their capability.

In addition to the wide range of logical connectives is their density in the passage. The insignificant result may have been caused by the fact that there were too many logical connectives in the passage. Altogether 20 logical connectives were added to the text. Just as under-signalling is detrimental to comprehension (see Hoey 1983), so it may be that over-signalling too may hinder comprehension. Morgan and Seller (1980) and Carrell (1982) argue that cohesive devices do not themselves contribute to coherence, but are interpreted through the readers' understanding of "content coherence." Coupland (1984) has talked of the "optimum level" of explicit signalling. Allison suggests that above the "optimum level", further cohesive ties will become irksome and counter-productive, while below it, the absence of explicit ties can interfere with comprehension (Allison 1991:378). If it is true that some logical connectives were actually not needed in the text, their presence would probably not aid comprehension, or might even impede comprehension. Thus, the number of logical connectives may have contributed to the insignificant result.

Finally, that logical connectives did not benefit comprehension may be due to the kind of reading strategy that students in the study employed. It has been argued that younger and less proficient students use fewer strategies and use them less effectively in their reading comprehension (Garner 1987, Nist and Mealey 1991, Padron and Waxman 1988). This is generally true of Hong Kong ESL readers, who pay too much attention to subordinate details, are very word-bound and inefficient in bottom-up processing, because they attend predominantly to the graphic form. The insignificant results of logical connectives could be explained by the fact that students did not use effective strategies to help comprehension. For example, they might not know how to make use of "selective attention", which is a kind of metacognitive strategy where students pay more heed to some features of the text, such as logical connectives in the study to enhance comprehension.

While headings and previews were found to benefit factual comprehension, none of the signals seemed to have an effect on inferential comprehension. Although some previous signalling studies have found significant results for inferencing, they have failed to provide a sound theoretical framework within which to explain the relationship between signals and inferencing. It seems reasonable to conclude from the study that signals do not aid inferencing. It could be argued that if signals highlight certain elements in the text and indicate the organization of the text, they cannot possibly help inferencing. Two further comments about signals and inferencing seem in order here. Firstly, signals tell us what is there in a text, whereas inference involves what is not there. Secondly, no attempt was made in the study to select signals, namely logical connectives which would focus attention on inferred information.

On the other hand, it is possible that signals could help inferencing, but then the so-called "floor effect" of the inferencing questions contributed to the insignificant result. What the floor effect means is that the test was not able to measure the entire range of achievement possible, where signals might have its influence. This is possible because the students in this study were not skilled readers and they might have found the inferential test so difficult that even the presence of signals did not help. As a result, they scored close to the minimum score, i.e. the "floor" of the test. This floor effect could be observed from the descriptive statistics. The mean score for inferential comprehension was far below 50% of the maximum mark. The standard deviation was about 1.95, indicating a small dispersion among the scores. This explanation, however, could be rejected on the grounds that if the hypothesis about signals was correct, even poor inferencing skills should be helped by signals.

As inferencing is an integral part of text comprehension, it would be useful to find out why students failed to infer. One reason why students did badly in inferencing is that although inferencing skill is not always difficult, it is frequently difficult in expository texts. According to O'Malley and Uhl Chamot (1990), inferencing is a kind of "cognitive strategy" which engages the readers in "interacting with the material to be learned, manipulating the material mentally or physically, or applying a special technique to a learning task" (O'Malley and Uhl Chamot 1990:138). In other words, inferencing as a cognitive strategy involves a number of cognitive skills, such as reasoning, logical thinking, and problem-solving. Another problem about inferencing is also that the existing curriculum in Hong Kong is not conducive to developing cognitive skills in students. Students are simply not sufficiently engaged in learning tasks where their cognitive thinking is encouraged and developed. As a result, a comprehension test that requires inferencing turns out to be very difficult for Hong Kong students. Moreover, it is a rather unfamiliar test format to them as the existing English language public examination for this level of students puts very little emphasis on the skill of inferencing.

A word of caution is necessary, however, in terms of generalizing the results externally. First of all, since the 285 subjects in the study are all female S4 students from three well-established English-medium schools in Hong Kong, identified as less skilled readers, the results of the study would thus be generalizable only to students of the same sex, of similar standards, backgrounds and school experiences. Secondly, only one passage was tested in the study, namely a non-technical prose passage about 500 words long, belonging to the text type "collection of descriptions". Hence, the results are not generalizable to other text types, texts of different lengths, and technical texts. Last but not least, since the study investigates headings, previews, and logical connectives only, the results could not be generalized to other signal types.

Nonetheless, the present research has demonstrated the value of including headings and previews in expository texts, which has significant implications for the design of reading materials for ESL readers. The inclusion of headings and previews in textbooks on content areas such as History and Geography would be particularly useful in helping students understand content knowledge taught through the medium of English. Although logical connectives in the study do not seem to help comprehension, they are not yet a closed issue. It would be interesting to find out if logical connectives would aid comprehension in other text situations, and how the possible inclusion of logical connectives could help ESL readers' comprehension. Given the tendency among ESL learners to misuse and overuse logical connectives in writing, there is a great need to investigate the role of logical connectives in discourse, with a view to helping ESL learners with proper processing of logical connectives in text comprehension, and hence appropriate use of connectives in writing. On the other hand, further research is needed to investigate the role of signalling in inferencing, which is found to be totally inadequate and ineffective among ESL readers in the study. Also the fact that students are weak in inferencing is certainly a sound justification for more questions of this nature to be included in reading comprehension tests, not less. The present imbalance in the comprehension question types found in reading tests in Hong Kong needs to be addressed.

Although headings and previews appear to aid factual comprehension, no conclusive evidence can yet be drawn about the whole issue of signalling and ESL reading comprehension. It is hoped that subsequent research on the area of signalling and ESL text comprehension will yield more insights into the role of signals in discourse, so that useful information can be yielded to help materials designers with the writing of texts for ESL readers, and to help teachers with more effective teaching of reading.

References

Afflerbach, P., and Johnston, P. (1984). Research methodology on the use of verbal reports in reading research. *Journal of Reading Behavior*, 16:307-321.

Allison, D. (1991). Textual explicitness and pragmatic inferencing: The case of 'hypothetical-real' contrasts in written instructional scientific discourse in English. *Journal of Pragmatics*, 15:373-393.

Britton, B.K., Glynn, S.M., Meyer, B.J.F. & Penland, M.J. (1982). Effects of text structure on use of cognitive capacity during reading. *Journal of Educational Psychology*, 74(1):51-61.

Carrell, P.L. (1982). Cohesion is not coherence. *TESOL Quarterly*, 16(4):479-488.

Cohen, A., Glasman, H., Rosenbaum-Cohen, P.R., Ferrata, J. & Fine, J. (1979). Reading English for specialized purposes: Discourse analysis and the use of student informants. *TESOL Quarterly*, 13(4):551-564.

Coupland, J. (1984). Writing texts: Cohesion in scientific and technical writing. In R. Williams, J. Swales & J. Kirkman (eds.), *Common ground: Shared interests in ESP and Communication Studies*, ELT Documents 117 (pp. 55-62). Oxford: Pergamon.

Crewe, W.J. (1990). The Illogic of logical connectives. *ELT Journal*, 44(4):316-325.

Dubin, F. & Olshtain, T. (1980). The interface of reading and writing. *TESOL Quarterly*, 14(3):353-363.

Freedle, R.O. & Hale, G. (1979). Acquisition of new comprehensive schemata for expository prose by transfer of a narrative schema. In R.O. Freedle (ed.), *New directions in discourse processing*. Norwood, NJ: Ablex.

Garner, R. (1987). *Metacognition and reading comprehension*. Norwood NJ: Ablex.

Geva, E. (1992). The role of conjunctions in L2 text comprehension. *TESOL Quarterly*, 26(4):731-747.

Glover, J.A., Dinnel, D.L., Halpain, D.R., McKee, T.K., Corkill, A.J. & Wise, S.L. (1988). Effects of across-chapter signals on recall of text. *Journal of Educational Psychology*, 80(1):3-15.

Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly*, 25(3):375-406.

Graesser, A.C. (1981). *Prose comprehension beyond the word*. New York: Springer.

Graesser, A.C., Hauft-Smith, K., Cohen, A.D. & Pyles, L.D. (1980). Advanced outlines, familiarity, and text genre on retention of prose. *Journal of Experimental Education*, 48:281-290.

Halliday, M.A.K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Hartley, J. & Trueman, M. (1983). The effects of headings in text on recall, search, and retrieval. *British Journal of Educational Psychology*, 53:205-214.

Hoey, M. (1983). *On the Surface of Discourse*. London: George Allen & Unwin.

Kintsch, W. & Yarbrough, J.C. (1982). Role of rhetorical structure in text comprehension. *Journal of Educational Psychology*, 74(6):828-834.

Klare, G.R., Shuford, E.H. & Nichols, W.H. (1958). The relation of format organization to learning. *Educational Research Bulletin*, 37:39-45.

Lee, I.K.B. (1993) Effects of signalling on ESL reading comprehension. Unpublished M.A. dissertation. University of Surrey.

Loman, N.L. & Mayer, R.E. (1983). Signalling techniques that increase the understandability of expository prose. *Journal of Educational Psychology*, 75(3):402-413.

Meyer, B.J.F. (1975). *The organization of prose and its effects on memory*. Amsterdam: North Holland.

Meyer, B.J.F. (1984a). Organizational aspects of text: Effects on reading comprehension and applications for the classroom. In J. Flood (ed.), *Promoting reading comprehension*, Newark: International Reading Association.

Meyer, B.J.F. (1984b). Effects of discourse type on recall. *American Educational Research Journal*, 21(1):121-143.

Meyer, B.J.F., Brandt, D.M. & Bluth, G.J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly*, 16(1):72-103.

Morgan, J.L. & Seller, M.B. (1980). Discourse and linguistic theory. In R.J. Spiro, B.C. Bruce & W.F. Brewer (eds.), *Theoretical issues in reading comprehension*. London: Lawrence Erlbaum.

Nist, S. & Mealey, D. (1991). Teacher-directed comprehension strategies. In Flippo & D. Caverly D (eds.) *Teaching reading and study strategies (pp. 42-85)*. Newark, DE: International Reading Association.

O'Malley, J.M. & Uhl Chamot. A. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

Padron, Y. & Waxman, H. (1988). The effect of ESL students' perceptions of their cognitive strategies on reading achievement. *TESOL Quarterly*, 22(1):146-150.

Rumelhart, D.E. (1975). Notes on a schema for stories. In D.G. Bobrow & A.M. Collins (eds.), *Representation and understanding: studies in cognitive science*. New York: Academic Press.

Spyridakis, J.H. & Standal, T.C. (1986). Headings, previews, logical connectives: Effects on reading comprehension. *Journal of Technical Writing and Communication*, 16(4):343-354.

Spyridakis, J.H. & Standal, T.C. (1987). Signals in expository prose: Effects on reading comprehension. *Reading Research Quarterly*, 22(3):285-298.

Spyridakis, J.H. (1989a). Signalling effects: A review of the research - Part I. *Journal of Technical Writing and Communication*, 19(3):227-240.

Spyridakis, J.H. (1989b). Signalling effects: Increased content retention and new answers - Part II. *Journal of Technical and Communication*, 19(4):395-415.

Thorndyke, P.W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive·Psychology*, 9:77-110.

Walker, C.H. & Meyer, B.J.F. (1980). Integrating different types of information in text. *Journal of Verbal Learning and Verbal Behavior*, 19:263-275.

Appendix 1

General Aviation

(Original Passage)

The modern traveller usually thinks of aviation in terms of airline transport, and perhaps of military applications. But the numerous uses of planes that come under the heading "general aviation" are often less known.

One of the major roles of general aviation is in business. In fact, business flying has become one of the world's fastest growing industries in recent years. Forecasters say that business flying will continue to grow despite the energy and fuel shortages in many areas. In short-distance travel, small aircraft use less fuel than big airliners. And as airlines continue to reduce the number of flights to smaller communities because of light and uneconomic loads, businessmen will continue to turn to company planes to get them to the right place at the right time. In many countries, business planes are no longer used only for top company officials. They now transport managers, engineers, accountants and other specialists who are required to travel a lot.

Busy people are also becoming used to the idea of hiring planes or taking small aircraft to get them to their destination quickly. In fact, such services frequently connect with the major airlines. The result is a more complete network of air communications similar to that which developed in earlier years in land and sea transport.

In the broad category of special purpose flying, there are many uses of general aviation. In the field of air cargo, aircraft pick up and drop off loads of cattle, drilling equipment, road-building machinery and life-saving supplies. Various sizes and designs of "general aviation" aircraft are used throughout the world in such tasks as hurricane hunting, the location of dangerous icebergs and forest fire fighting. In recent years, there has been increased interest in flying sports such as gliding, manpowered flight and hang-gliding.

But these spectacular uses are only part of the story. General aviation in all of its forms now contributes thousands of interesting and highly skilled jobs to the economies of the nations, with a growing number of small airports providing new trade and communication networks.

The recent development of the short take-off and landing aircraft (STOL) has boomed in general aviation. It paves the way to flights to many areas that have only short runways. The small community with an industrial airport can now compete on a more equal basis with big cities in attracting industries. In fact, the growth of general aviation will lead to decentralization in industries. Companies can be attracted to smaller population centres by their closeness to raw materials, by their lower land and operating costs, and by better living conditions for employees. General aviation makes it more convenient for remote communities to reach larger market centres.

If the past is any guide, man will continue to find new special uses for the aeroplane.

(From *Mastering English Precis and Language Practice*)

Appendix 2

General Aviation

(With Signals)

The modern traveller usually thinks of aviation in terms of airline transport, and perhaps of military applications. But the numerous uses of planes that come under the heading "general aviation" are often less known. *The various uses of general aviation will be discussed in this passage*.

Business Flying

One of the major roles of general aviation is in business. Business flying has become one of the world's fastest growing industries in recent years. In fact, forecasters say that business flying will continue to grow despite the energy and fuel shortages in many areas. One reason is that in short-distance travel, small aircraft use less fuel than big airliners. Also, as airlines continue to reduce the number of flights to smaller communities because of light and uneconomic loads, businessmen will continue to turn to company planes to get them to the right place at the right time. As a result, in many countries, business planes are no longer used only for top company officials. They now transport managers, engineers, accountants and other specialists who are required to travel a lot.

Busy People

Moreover, *general aviation offers a useful service to busy people*, who are becoming used to the idea of hiring planes or taking small aircraft in order to get them to their destination quickly. In fact, such services frequently connect with the major airlines. The result is a more complete network of air communications similar to that which developed in earlier years in land and sea transport.

Special Purpose Flying

In the broad category of special purpose flying, there are many uses of general aviation. For instance, in the field of air cargo, aircraft pick up and drop off loads of cattle, drilling equipment, road-building machinery and life-saving supplies. Besides, various sizes and designs of special purpose "general aviation" aircraft are used throughout the world in such tasks as hurricane hunting, the location of dangerous icebergs and forest fire fighting. Furthermore, in recent years, there has been increased interest in flying sports such as gliding, manpowered flight and hang-gliding.

General Aviation and Economy

General aviation has some bearing on the economy. General aviation in all of its forms now not only contributes thousands of interesting and highly skilled jobs to the economies of the nations, but it also results in a growing number of small airports providing new trade and communication networks.

General Aviation and Industries

General aviation contributes to the industries in a significant way. The recent development of the short take-off and landing aircraft (STOL) has boomed in general aviation. It thus paves the way to flights to many areas that have only short runways. Because of this the small community with an industrial airport can now compete on a more equal basis with big cities in attracting industries, which will therefore lead to decentralization in industries. In addition, companies can be attracted to smaller population centres by their closeness to raw materials, by their lower land and operating costs, and by better living conditions for employees. Thus, general aviation makes it more convenient for remote communities to reach larger market centres.

New uses

Last but not least, if the past is any guide, man will continue to find new special uses for the aeroplane.

Note: Headings are in **boldface** and underlined.
Previews are *italicized* and underlined (in normal face type and not underlined in actual tests).
Logical connectives are underlined (but not underlined in actual tests).

Appendix 3

Comprehension Questions

1. List the 5 uses of general aviation mentioned in the passage.

   a.

   b.

   c.

   d.

   e.

2. What are the advantages of small aircraft over airlines for business people?

3. Why is there now such a wide range of aircraft designs?

4. If the general aviation industry declined, what would the effect be on small communities?

5. What is the likely future of general aviation?

6. What do you understand by the term "general aviation"?

Appendix 4

Table 1

Mean Scores and Standard Deviations

	Mean Score	s.d.	No. Subjects
No Signals	4.20 F 3.27 I	2.78 F 2.10 I	30
Headings Only	6.84 F 3.36 I	3.37 F 1.87 I	22
Previews Only	6.50 F 3.02 I	3.09 F 1.72 I	27
Logical Connectives Only	4.85 F 2.55 I	2.44 F 1.98 I	30
Headings & Previews	7.70 F 3.38 I	3.54 F 2.26 I	30
Headings & Logical Connectives	6.30 F 3.03 I	3.30 F 1.49 I	35
Previews & Logical Connectives	5.93 F 2.78 I	3.33 F 2.39 I	30
Headings, Previews & Logical Connectives	8.09 F 3.39 I	2.46 F 1.87 I	28

F = factual comprehension
I = inferential comprehension

Table 2

Summary of Results for Individual Signals (p value)

	ANOVA		
Hypothesis	**Factual**	**Inferential**	**MANOVA**
Headings Only	.0039*	.7369	.0163*
Previews Only	.0120*	.6203	.0404*
Logical Connectives Only	.5460	.3873	.5812

* = p < .05

Table 3

Summary of Results for Two Signals in Combination (p value)

	One-Way ANOVA		
Hypothesis	**Factual**	**Inferential**	**MANOVA**
Headings & Previews	.0001*	.6231	.0003*
Headings & Logical Connectives	.0111*	.7661	.0403*
Previews & Logical Connectives	.0444*	.5147	.1175

* = p < .05

Table 4

Summary of Results for Three Signals in Combination (p value)

Hypothesis	One-Way ANOVA Factual	 Inferential	MANOVA
Headings, Previews & Logical Connectives	.0001*	.6579	.0001*

* = p < .05

Table 5

Results of Three-way ANOVA (p value)

Main / Interaction Effects	F	I	MANOVA
Headings Only	.0001*	.1342	.0001*
Previews Only	.0002*	.7207	.0011*
Headings x Previews	.6491	.6993	.8324
Logical Connectives Only	.9663	.2167	.4670
Headings x Logical Connectives	.8843	.5442	.8212
Previews x Logical Connectives	.8592	.4219	.7121
Headings x Previews x Logical Connectives	.1846	.8945	.4094

F = factual comprehension
I = inferential comprehension
* = p < .05

THE VARIETY AND EFFECTIVENESS OF STRATEGIES EMPLOYED IN VOCABULARY EXPLANATIONS IN EFL CLASSROOMS IN HONG KONG

Annie Lee On-lai

Introduction

This study examines the effectiveness of two factors in vocabulary teaching; firstly, the presence or absence of a text and secondly the use of a variety of explanation types (hereafter referred to as multi-type explanations) or a single explanation type (hereafter referred to as mono-type explanations).

Vocabulary teaching is one aspect of language teaching that has not been given the attention it deserves until recently. Wilkins (1972:111) emphasizes the need for vocabulary teaching. He maintains that '(W)ithout grammar very little can be conveyed, without vocabulary nothing can be conveyed.' Yet without a deeper understanding of how vocabulary is taught in the classroom and which methods of teaching are more effective for learners, the teaching of vocabulary may not achieve the desired effects in the classroom where maximum teacher input and assistance are available.

A number of studies (e.g. Nation 1982, 1983, Carter 1985, 1987, and see also Bird, this volume) has been conducted on vocabulary acquisition. Nation and Coady (1988) suggest that 'context' for vocabulary learning can be viewed as both context within a text and the reader's background knowledge of the subject matter of a given text. In discussing the effect of context on vocabulary learning, they comment that 'studies on learning words from context have not shown the large amounts of learning we might expect, considering the rates at which first-language learners seem to increase their vocabulary.' (ibid.:103)

Nation (1990:1-2) further states that 'many teachers too quickly dismiss the approach of getting learners to study lists of words out of context', which can be a very effective way of acquiring vocabulary within a short period of time. On the other hand, Carter and McCarthy (1988:109) make the following observation about the importance of context in vocabulary learning, '(T)he general conclusion to be drawn from research is that learning vocabulary through context must be the major way of increasing vocabulary knowledge.' How far context (context within a text and background knowledge of the reader) may assist vocabulary explanation in the classroom thus forms one of the two factors to be examined in this study.

To define word meaning, Nation (1990) recommends four types of definition: definition by demonstration, by abstraction, through context and by translation. While each of these types of definition can be adopted to explain vocabulary items, Nation asserts that it is best to combine two or more types of definition. The second factor examined in this study, therefore, is whether a single or a variety of explanation types better assist learners to abstract word meaning.

Despite the extensive research into vocabulary acquisition, classroom research on vocabulary explanation has not been widely conducted. This research therefore aims to investigate how and to what extent the two factors mentioned above affect vocabulary learning in the classroom. Statistical data and classroom discourse data will be examined.

Identification of Vocabulary Teaching Strategies

Before investigating the effectiveness of the two factors in vocabulary explanation, vocabulary teaching strategies in EFL classrooms were identified. Three S3 (Grade 9) and one S4 (Grade 10) reading comprehension classes from two different schools taught by four different teachers were observed and video-taped. All four teachers had more than five years of EFL experience, and two of them had majored in English in their university studies. The strategies used by these teachers were then classified according to Nation's (1990) vocabulary definitions with modifications made by the researcher. They are listed as follows:

A. Non-verbal

1. Using objects

Teachers bring into the class real objects or point to real objects within viewing distance of students.

2. Using blackboard drawings

Teachers use blackboard drawings for objects that cannot be brought into the classroom to illustrate what they are or what they look like.

3. Using pictures

Most course books are richly illustrated with pictures and these provide a useful and direct source of explanation. Teachers may also bring into the class pictures to illustrate meanings of words.

4. Demonstrations

Most action words and many situations can be more clearly explained by demonstration in class. For example, in explaining the word "rickshaw", the teacher got a student to demonstrate how a rickshaw is pulled.

5. Using gestures

a. Paralinguistic gestures help to convey meaning but are not essential to the explanation of word meaning.

b. Non-paralinguistic gestures may be essential to the explanation of word meaning since, without them, the meaning may not be effectively communicated, e.g. in explaining the word "terrified", the teacher used gesture to convey that she was terrified.

B. Verbal

1. Using synonyms

Although no two English words share exactly the same meaning, sometimes words can be explained by synonyms to illustrate one dimension of the meaning. For example, "rapidly" was explained by a teacher as "quickly".

2. Paraphrasing

When no synonyms can conveniently replace the meaning of the word, the meaning is explained in one or more phrases/ clauses.

3. Exemplifications

Teachers provide examples in the hope that students can deduce the meaning of the vocabulary item from the example.

4. Dictionary definitions

These are like paraphrases of word meaning but exact details are provided.

5. Using affixes / word-roots

Teachers use affixes and word-roots to help students get the basic meaning of a vocabulary item from which the exact word meaning is developed.

6. L1 Explanation

a. Teachers offer explanations in Cantonese (no such example was obtained from all the transcribed lessons).

b. Teachers solicit L1 explanation

If teachers expect students to know the word before or after an explanation is given, they may solicit a Cantonese explanation from the students as confirmation.

Design of the Study

An experiment was then designed to observe the effectiveness of multi-type and mono-type explanations with and without a text to establish how much, if any, these factors affect vocabulary learning in class. The 2 factors studied were manipulated so that the 4 groups were each given a different treatment as follows:

Group 1	Subjects were given a text and multi-type explanations of 10 target vocabulary items selected from a text.
Group 2	Subjects were given the same text and mono-type explanations of the same vocabulary items.
Group 3	Subjects were given multi-type explanations of the same 10 target vocabulary items out of context, that is, without a text.
Group 4	Subjects were given mono-type explanations of the same vocabulary items out of context, that is, without a text.

Two upper stream S3 (Grade 9) classes from a band 4 - 5 girls' school in the Hong Kong district were divided into four groups of subjects for the experiment. (Band 1 is the top of academic banding and Band 5 the bottom.) These groups, of 20 - 22 students each, were taught by the same teacher to control the teacher variable that might affect class input and thus learning outcome. The teacher, an English major university graduate with fifteen years of EFL experience, had not taught any of the subjects before. Thus her ways of teaching, if foreign to the subjects, were foreign to all subjects. While the teachers of the four classes video-taped (from which vocabulary teaching strategies were identified) were not conscious of the objectives of the study; the teacher conducting the experiment, being the researcher herself, was consciously aware of the study objectives as well as the importance of manipulating the four types of treatment.

The experiment adopted a pre-test -> treatment -> post-test design. Before effecting different treatments to the different groups, all groups were given a pre-test to establish the fact that the majority of the ten target vocabulary items (qualifications, approach, thoroughly, initiative, ambitions, pamphlets, vacancy, hunt, prospects, confident) were not previously known. This pre-test required the subjects to explain 19 vocabulary items (ten of which were the target items) in English and/or Chinese so that pre-treatment identification of the target vocabulary was avoided.

After the pre-test, each group was given a different treatment as mentioned above. Fifteen minutes after each treatment, the subjects were given 3 post-tests. Post-test 1 was a replication of the pre-test while post-test 2 was a summary cloze based on the given text and required subjects to fill in the blanks with the target vocabulary items out of a pool of 20 given words. Post-test 3, made up of 10 isolated sentences, required subjects to fill in the blanks with the same target vocabulary items out of the same word pool.

After the treatment and post-test sessions, 3 students from each group, selected at random, were invited for a video-taping and interview session to provide feedback on their learning during the treatment lessons and what they perceive were clear and effective vocabulary explanations.

Analysis of Findings

A comparison of the performance of the four groups in the pre-test and post-tests indicates whether teaching the subjects with or without a text using either multi-type or mono-type explanations has led to significant differences in vocabulary learning. Since the pre-test and post-test 1 were essentially the same test, a comparison of the results of the two tests illustrates the effects of each treatment as far as meaning recall of the target vocabulary items was concerned. Post-test 2, being a summary cloze of the text given to groups 1 and 2 but not groups 3 and 4, tested only how groups 1 and 2 applied the target vocabulary items to a familiar context (the context in which the vocabulary items were learnt). To groups 3 and 4 (the groups not given a text), post-test 2 was a test of vocabulary application to an unfamiliar context. Post-test 3, being a test of vocabulary application to contexts different from that of the text, was written in contexts unfamiliar to all four groups.

Table 1 illustrates the performance of the four groups of subjects in the pre-test and post-tests regarding the 10 target vocabulary items.

Presence or Absence of a Text

Since post-test 2 was not a familiar context to all four groups, only the results of the four groups in post-test 1 and post-test 3 could be compared and studied. Taking these post-tests to correspond to two dependent variables, a multivariate analysis of variance was performed (see Tables 2 and 6). An overall comparison of the scores of groups 1 and 2 (groups with text) and those of groups 3 and 4 (groups without text) indicates that the presence of a text does have a significantly positive effect on test results, as illustrated by Table 2.

Table 1

Overall Performance of the 4 Groups in the 10 Target Vocabulary Items

	n=20				n=21				n=21				n=22			
Vocabulary	Group 1				Group 2				Group 3				Group 4			
	multi, + text				mono, + text				multi, - text				mono, - text			
	Pr	Po1	Po2	Po3	Pr	Po1	Po2	Po3	Pr	Po1	Po2	Po3	Pr	Po1	Po2	Po3
qualifications	1	8	0	9	0	12	1	11	0	17	2	8	0	7	2	0
approach	0	8	9	3	1	11	11	5	0	6	2	4	0	1	0	0
thoroughly	0	9	5	3	0	14	4	2	0	1	3	0	0	3	1	1
initiative	0	1	0	0	0	9	1	4	0	1	0	0	0	4	1	0
ambitions	0	5	3	1	0	9	0	1	0	10	2	1	0	1	1	1
pamphlets	2	12	11	13	0	19	13	14	0	5	1	4	0	7	0	5
vacancy	0	14	4	7	1	9	4	7	0	1	0	3	0	1	1	2
hunt	7	15	14	3	9	14	8	0	12	16	0	0	3	9	1	1
prospects	0	7	5	2	0	6	3	1	0	6	2	0	0	1	2	3
confident	4	15	11	6	5	19	4	2	0	14	3	3	0	8	3	2
Total	14	94	62	47	16	122	49	43	12	77	14	23	3	42	12	15

Pr = pre-test, Po1 = post-test 1, po2 = post-test 2, po3 = post-test 3

Table 2 shows that the null hypothesis for factor 1 (the presence or absence of a text) is rejected, indicating that the text factor does have a significant effect on the performance of the subjects. Table 3, which shows the score differences across the ten vocabulary items, better illustrates the actual performance of the subjects on individual items.

Table 2

Multivariate Analysis of Variance for

the Hypothesis of no Overall Factor 1 (presence or absence of a text) Effect

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilks' Lambda	0.525	17.399	4	77	0.0001
Phillai's Trace	0.475	17.399	4	77	0.0001
Hotelling-Lawley Trace	0.904	17.399	4	77	0.0001
Roy's Greatest Root	0.904	17.399	4	77	0.0001

Table 3

Score Differences of Groups 1 and 2 (with text) Versus Groups 3 and 4

(without text) in Post-test 1 and Post-test 3

Target Vocabulary	Grps 1 & 2 Results vs Grps 3 & 4 Results in Post-test 1	Grps 1 & 2 Results vs Grps 3 & 4 Results in Post-test 3
qualifications	-	+
approach	+	+
thoroughly	+	+
initiative	+	0
ambitions	-	0
pamphlets	+	+
vacancy	+	+
hunt	-	+
prospects	+	0
confident	+	+

+: Groups 1 and 2 obtained higher scores than groups 3 and 4
-: Groups 1 and 2 obtained lower scores than groups 3 and 4
0: No score difference between groups 1 & 2 and groups 3 & 4

In both post-tests 1 and 3, groups 1 and 2 performed better than groups 3 and 4 on seven out of the ten target vocabulary items. While groups 1 and 2 obtained lower scores than groups 3 and 4 in three vocabulary items in post-test 1, they obtained the same scores as the two latter groups in three items in post-test 3. As groups 1 and 2 were given multi-type and mono-type explanations with a text and groups 3 and 4 were given the same explanations without a text, the use of a text does significantly affect the learning outcome of vocabulary explanations. The text not only provides a first-hand context of vocabulary items through which vocabulary meaning can be guessed, but also serves as a concrete example of how these items are used and collocated with other words.

The findings are further supported by student comments in the post-experiment interview sessions. The majority of subjects from groups 3 and 4 (without text) stated that they could learn vocabulary more effectively with a text because they could guess the meaning of the words as well as examine how the words were used in relation to other words. Subjects in group 4 maintained that the presence of a text was essential, and that without one, vocabulary explanation must be made with examples so that word meaning could be clearly comprehended. Otherwise, they added, they often felt that the meaning of one word was very similar to the meaning of another word and only L1 explanation could help them to distinguish one from the other. The positive effect of using a text in vocabulary explanations is thus further reinforced.

Among the three target vocabulary items for which groups 1 and 2 obtained lower scores than groups 3 and 4 in post-test 1, two of them, 'qualification' and 'hunt' can be explained when interviews with students and the other types of scores for the four groups were examined.

Although very few subjects (in fact only 1) among the 84 subjects got the meaning of 'qualifications' correct in the pre-test, a significant number of them got it correct in post-tests 1 and 3. When subjects were interviewed, they said that they were able to remember the meaning of the word because once they heard the pronunciation of the word, they immediately recognised it to be the same word as the Cantonese transliteration of the word, "quali", which was often used in Cantonese conversations. In other words, before the teacher's explanation, they in fact knew the word by sound but not by form. The fact that they could remember the meaning of the word has to do with the matching of the sound to the form of the word rather than to do with the teacher's explanation.

As for 'hunt', the seemingly higher scores of groups 3 and 4 in post-test 1 have to be considered with the scores of the four groups in the pre-test. The score differences of the four groups between the pre-test and post-test 1 are as follows:

Table 4

Scores Differences Between Pre-test and Post-test 1 Regarding 'Hunt'

	Group 1	Group 2	Group 3	Group 4
Scores Differences	8	5	4	6

When the differences of the scores are considered instead of the raw scores of post-test 1, groups 1 and 2 performed slightly better than groups 3 and 4. Hence, even with a comparatively more familiar item such as 'hunt', the use of a text better helped students abstract the meaning of the word.

Use of Multi-type or Mono-type Explanations

A comparison of the scores between groups 1 and 2 (the multi- type and mono-type groups with text) as well as between groups 3 and 4 (the multi-type and mono-type groups without text) shows whether multi-type explanations, as recommended by Nation (1990), work better than mono-type explanations. Table 5 summarizes the overall effect of this second factor studied in the experiment.

Table 5

Multivariate Analysis of Variance for the Hypothesis of no Overall Factor 2 (use of multi-type or mono-type explanations) Effect

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilks' Lambda	0.935	1.334	4	77	0.2649
Pillai's Trace	0.0648	1.334	4	77	0.2649
Hotelling-Lawley Trace	0.0693	1.334	4	77	0.2649
Roy's Greatest Root	0.0693	1.334	4	77	0.2649

The null hypothesis for the second factor has not been rejected, indicating that the use of multi-type or mono-type explanations in this instance has no significantly different effect on the subjects' scores. The following table compares whether the scores obtained by the multi-type groups are higher or lower than the mono-type groups.

Table 6a

Score Differences Between Group 1 and Group 2 in Post-tests 1 - 3

Vocabulary Items	Post-test 1	Post-test 2	Post-test 3
qualifications	-	-	-
approach	-	-	-
thoroughly	-	+	+
initiative	-	-	0
ambitions	-	+	0
pamphlets	-	-	-
vacancy	+	0	0
hunt	+	+	+
prospects	+	+	-
confident	-	+	+

-: Group 1 (the multi-type group) obtained lower scores than group 2 (the mono-type group)
+: group 1 (the multi-type group) obtained higher scores than group 1 (the mono-type group)
0: Group 1 (the multi-type group) and group 2 (the mono-type group) obtained equal scores

The score differences do not indicate that the multi-type groups (with and without text) necessarily obtain higher scores than the mono-type groups. In other words, multi-type explanations, whether given with or without a text, do not necessarily lead to better learning outcomes than mono-type explanations.

A careful study of the discourse data shows that the clarity and lucidity of vocabulary explanation, whether multi-type or mono-type, can be affected by a number of other factors. These are: (1) the level of learner involvement, (2) the repetition of vocabulary items, (3) elaboration of vocabulary explanations, (4) the level of abstraction of the vocabulary item and (5) the presence or absence of L1 equivalents.

Table 6b

Score Differences Between Groups 3 and 4 in Post-test 1 and 3

Vocabulary Items	Post-test 1	Post-test 3
qualifications	+	+
approach	+	+
thoroughly	-	-
initiative	-	0
ambitions	+	0
pamphlets	-	-
vacancy	0	-
hunt	+	-
prospects	+	-
confident	+	+

-: Group 3 (the multi-type group) obtained lower scores than group 4 (the mono-type group)
+: Group 3 (the multi-type group) obtained higher scores than group 4 (the mono-type group)
0: Group 3 (the multi-type group) and group 4 (the mono-type group) obtained equal scores

The Level of Learner Involvement

Since it has been established that the use of text has a significantly positive influence over vocabulary teaching and thus learning, results of the non-text groups are free from the effects of text use. The overall results of the three post-tests show that group 3 performed better than group 4. This could be due to two possible reasons. Firstly, the richer variety of explanations could have made meaning abstraction easier for group 3. Secondly, the vocabulary explanations of group 3, which make higher cognitive demands on subjects by involving them in producing examples or by involving them in processing word meaning through examples, better help them grasp word meaning. Two extracts of the discourse data taken from groups 3 and 4 in the explanation of 'ambitions' help illustrate this:

Discourse data - Experiment Group 3

G117 Ms. Lee: Alright? Er, another word. (writes on board)
G118 Ambitions.
G119 Ss : Ambitions.
G120 Ms. Lee: Ambitions.
G121 Ss : Ambitions.
G122 Ms. Lee: Now, ambitions are things you want to be or
G123 you want to do. Well, like for example, when
G124 I was a small girl, my ambition was to be,
G125 what? No, not a teacher, was to be an animal
G126 doctor, OK? Have you got any ambitions?
G127 (to a student) What is your ambition?
G128 S1 : Nurse.
G129 Ms. Lee: You want to be a nurse. (to another student)
G130 Yours? Yes, you? Yes, have you got any
G131 ambitions? (Ss laugh) Nothing? (to another
G132 student) You?
G133 S2 : A teacher.
G134 Ms. Lee: To be a teacher. OK? Ambitions, repeat,
G135 ambitions.
G136 Ss : Ambitions.

Discourse data - Experiment Group 4

H64 Ms. Lee: OK, another one. (writes on board)
H65 Ambitions.
H66 Ss : Ambitions.
H67 Ms. Lee: Ambitions.
H68 Ss : Ambitions.
H69 Ms. Lee: You know what ambitions are? Em ambitions
H70 are things you want to be or things you want
H71 to do. OK? Ambitions.
H72 Ss : Ambitions.
H73 Ms. Lee: Ambitions.
H74 Ss : Ambitions.

The fact that involving learners in processing word meaning is a better explanation strategy is also supported by Nation (1990) and Nattinger (1988). Nation (ibid:63-64) states that "(I)n order for learning to last, the learner must make an effort. The best way to make sure a learner forgets a word is for the teacher to present a short, clear explanation of the meaning and then pass on to the next piece of work.' It is only through active learner involvement that deep learning can be effected. Deep motivated learners, according to Biggs and Moore (1993:312), relate 'the content to personally meaningful contexts or to existing prior knowledge.'

A rich variety of explanation strategies that involve learners in mental processing has been noted when four EFL lessons were video-taped during which the range of vocabulary strategies were identified. These include: instructing students to repeat the new vocabulary item by helping them associate the sound with the form of the word, involving students in associating meaning with the word by saying the word and performing paralinguistic gestures at the same time to help them better remember the words, asking students to explain a word by means of demonstrations and examples, soliciting an L1 explanation of the word from students, conducting immediate tests which re-stimulate students to process the words learnt and lastly, repetition of the words learnt through subsequent tasks that engage students in further rethinking or reprocessing of the words.

The Repetition of Vocabulary Items

On examining subjects' answers in post-test 1 in which subjects were asked to put down the meanings of the vocabulary items in either English or Chinese, it was found that there are a number of instances where they put down the teacher's explanation against the wrong item. And out of the 10 target vocabulary items, those which have the fewest instances of the subjects misidentifying the meaning of the word are 'qualifications', 'initiative', 'vacancy', 'hunt', and 'confident'. Evidence suggests that at least three of these five items had been previously encountered by the subjects prior to the experiment.

Firstly, it has been stated above that the subjects admitted having heard of the Chinese transliteration of 'qualifications' in the post-experiment interview session. Secondly, the pre-test result of 'hunt' demonstrates that this is another familiar item. Finally, successful L1 translation solicitations for 'confident' in the experiment proves that this is also a familiar vocabulary item (the low incidence of mismatching for 'initiative' will be discussed below). It cannot be sheer co-incidence that a lower incidence of misidentification occurs among more familiar items.

For a learner to successfully acquire a vocabulary item, the same item should be repeated in the same lesson and in subsequent lessons as discussed in the sub-section on the level of learner involvement. Channell (1980) maintains that an important part of vocabulary acquisition course is its exercise so that the words will be repeated not only in the lesson they are taught but also in subsequent lessons.

A less familiar item can at most form a learner's passive vocabulary and only very familiar items become a learner's active vocabulary that can be encoded in the productive skills of writing and speaking.

Elaboration of Vocabulary Explanations

Despite the fact that clear and short explanations can easily be forgotten and multi-type explanations are usually richer, over-elaborate explanations may merely serve to confuse and distract. Thus with 'initiative', probably the most difficult item on the list, the mono-type groups may find a short paraphrase inadequate to deduce word meaning while the over-elaborate explanations that take up 28 and 27 transcribed lines of teacher talk in groups 1 and 3 respectively might have caused ambiguity and confusion and hence have defocused students' attention on the meaning of the word.

Discourse data - Experiment group 1

E321 Ss : Initiative.
E322 Ms. Lee: Initiative.
E323 Ss : Initiative.
E324 Ms. Lee: Initiative.
E325 Ss : Initiative.
E326 Ms. Lee: Now, initiative comes from the word initiate.
E327 To initiate means to start something. If you
E328 start something, then you initiate something.
E329 So if you show initiative in something, that
E330 means you don't have to wait for people to
E331 tell you what to do. You can start doing
E332 something on your own. That is to show er
E333 initiative in your work, OK? Now, for
E334 example, if you if you are a good student
E335 you should show initiative in learning. How
E336 do you show initiative in lear-learning? You
E337 don't have to wait for the teacher to tell you
E338 what to do and what to read. You can do
E339 things on your own, OK? You know when
E340 you should finish your homework, when you
E341 start your revision, you know when to go to
E342 the library and borrow books and learn more
E343 about a subject, alright? To show initiative,
E344 alright? You can do things on your own. You
E345 can start working on your own without waiting,
E346 without having to wait to be told what to do.
E347 OK, now let us look at the last paragraph.

As both the mono-type groups and multi-type groups had difficulty grasping the meaning of 'initiative', they probably singled this out as the most difficult item, thereby explaining the low incidence of misidentification with this item. Chaudron (1982:170) states that 'a major problem for the student may lie in the teacher's elaboration of vocabulary meanings through increased redundancy; the non-native listener may find it difficult to decode the exact message, because he cannot discern whether the same information has been provided redundantly or whether new information has been supplied.' The extent of elaboration in multi-type vocabulary explanations cannot be overlooked.

The Level of Abstraction of Vocabulary Items

Nation (1990) suggests that the learning burden of a word can be affected by the part of speech of a word. According to him, nouns and adjectives are easier to learn than verbs and adverbs. Yet the overall performance of the four groups illustrated in Table 1 shows that the most difficult item is 'initiative'. Although 'initiative' is a noun, the subjects did not find it any easier than 'thoroughly', an adverb.

One added feature that determines whether a word is difficult or not is its level of abstraction. The more abstract is the item, the more ambiguous is the meaning and thus more demanding is the task of meaning abstraction. As abstract concepts cannot be acquired from our senses, these are naturally more difficult to learners. Hence, whether to employ multi-type or mono-type explanations appropriately depends closely on how concrete or abstract an item is.

The Presence or Absence of L1 Equivalents

Another feature that explains whether a vocabulary item is easy or difficult to learners is whether there is an L1 equivalent for the item. Since L1 association greatly reduces the learning load and helps learners develop new concepts from known ones, an item with an L1 equivalent is undeniably easier for learners than those without. This being so, 'initiative' is a difficult item not only because it is a highly abstract item, but also because no L1 equivalent exists for the word.

Conclusion

The two factors examined in this study were whether the use of a text better assists vocabulary explanations in the classroom and whether multi-type or mono-type explanations better help learners abstract word meaning. The experimental results support the use of a text in vocabulary explanations but do not indicate whether multi-type or mono-type explanations work better at offering clear and lucid explanations of word meaning.

Multi-type or mono-type explanations cannot be assessed independently from a variety of other factors that affect learning: the level of learner involvement, the repetition of vocabulary items, the elaboration of vocabulary explanations, the level of abstraction of vocabulary items and the presence or absence of L1 equivalents. Indeed, as with classroom teaching in general, explanations cannot and should not be made without careful consideration of learners as an active participant, the inherent difficulty of the vocabulary to be taught and the quality of teacher input.

Direct interaction with learners in the classroom provides valuable on-the-spot feedback to teachers who can then make instantaneous decisions on whether an explanation is comprehended or whether further elaborations, repairs or repetitions are necessary. It is such dynamics of classroom interaction that make classroom research a frutiful and intriguing area of investigation.

References

Biggs, J.B. & Moore, J. (1993). *Process of learning*. Sydney: Prentice Hall.

Carter, R. & McCarthy, M. (1988). *Vocabulary and language Teaching*. Harlow: Longman.

Channell, J. (1980). Applying semantic theory to vocabulary teaching. *English Language Teaching*, 35:115-122.

Chaudron, C. (1982). Vocabulary elaboration in teacher's speech to L2 learners. *Studies in Second Language Acquisition*. 4(2):170-180.

Nation: (1990). *Teaching and learning vocabulary*. Boston, MA: Heinle and Heinle.

Nation & Coady, J. (1988). The role of dictionaries in language learning. In R. Carter & M. McCarthy (eds.) *Vocabulary and language teaching*. Harlow: Longman.

Nattinger, J. (1988). Psycholinguistic considerations in the study of L2 vocabulary acquisition. In R. Carter & M. McCarthy (eds.) *Vocabulary and language teaching*. Harlow: Longman.

Wilkins, D.A. (1972). *Linguistics and language teaching*. London: Edward Arnold.

PART FOUR

RESOURCES FOR TEACHERS AND LEARNERS

INTRODUCTION

Under the sub-theme "Resources for Teachers and Learners", the following issues were specified for discussion at ILEC 1993:

1. resourcing for students with different levels of learning and performance;
2. constructing banks of texts, tasks and reference materials for learning;
3. constructing a bank of assessment tasks;
4. reassessing the role of textbooks;
5. using resources to develop independent learning;
6. resourcing for task-centred language learning;
7. using television as a learning resource;
8. using interactive computer networks for teacher and learner support.

The "resources" sub-theme also attracted discussions on a number of original topics that appealed to the programme organisers. On the other hand, some suggested issues (such as the construction of "banks" of materials or tests) were not, in the event, widely addressed.

The eventual selection of papers for the Conference, and more strictly for this volume, was governed by the view that papers on this sub-theme should be characterised by at least two of the following traits:

- relevance to learning (given the overall conference theme);

- either an evident link to one of the suggested issues for discussion, or an original and pertinent view of what counts as a "resource";

- an agenda that was not promotional but primarily research-based.

Being aware of other conferences and published collections relating to some aspects of "resources", notably self-access learning and computers in language learning, the editors were particularly stringent in assessing the relevance of papers in these areas to the specified concerns of ILEC 1993. An effort was made to include school-based work as well as work arising from tertiary education, which was incidentally strongly represented at the Conference.

In the first paper in this section, *Vivienne Yu, Emily Chiu, Winnie Siu and Rachel Yau* describe a territory-wide questionnaire survey of the use of extensive reading materials in Hong Kong primary schools. Findings from 294 schools suggest that teachers see students as having considerable problems with reading and that extensive reading is widely considered as useful; actual teaching practices may, however, not adequately encourage extensive reading as little class time is devoted to it. It is thought that a proposed Primary English Extensive Reading Scheme given the necessary encouragement and resource materials could well influence the use of teaching time for reading activities. 120 schools have indicated an interest in participating in the new scheme.

Reading is important for many reasons, not least among which is the impact it can have on vocabulary size and range. Language, as Halliday points out, is a resource for making meaning. Vocabulary is obviously a core component of this resource, and is not only a product of a language learning and using process but also a crucial resource for further learning. *Norman Bird* reports on the second stage of his research into the knowledge of English vocabulary demonstrated by native and non-native speakers of different age groups and levels of educational attainment in their performance on vocabulary recognition tests described in his paper. Among other findings, Bird's results show that vocabulary size of non-native speakers (and to some extent of native speakers, whose vocabulary size is predictably greater) increases both with age and with educational level, thereby quantifying the association between education, life experience and vocabulary size.

A core element in traditional educational resources is "the textbook", but limitations of textbooks in contemporary education have also been much discussed. *Bob Adamson and John C.K. Lee* present "textbook resources" as a term that includes both a core text and associated written or audio-visual learning materials, and assess the future role of textbook resources as Hong Kong moves towards the proposed "Target-Oriented Curriculum (TOC)" which they associate with a "Progressivist" approach to education. They conclude that Progressivist textbook resources will be needed in order to achieve a balance, or a compromise, between ideal expectations and practical realism in education.

Balancing the desirable and the attainable is also a theme in the paper by *Valerie Pickard, Kenneth Chan and Janice Tibbetts* on "Concordancing in the Classroom". This account presents one relatively novel alternative (or complement) to the traditional textbook. Describing concordancers as "simple computer programs which can quickly analyze electronic texts to find any occurrence of a given word, part of a word or phrase and display it within its immediate context", the paper shows how concordancers can help teachers and students to explore and by implication extend their vocabulary knowledge. The paper goes on to discuss criteria for selecting a corpus/corpora, i.e. set(s) of texts for analysis for Hong Kong schools. The final section takes a practical and probing look at constraints on the use of computers and concordancing in schools, and also illustrates its potential for motivating secondary students.

Educational television offers another potential language teaching and learning resource. In one of several conference contributions that addressed the theme of authenticity of materials and activities, *Arthur McNeill* argues for the potential value of unscripted and semi-scripted programmes that feature interviews with local Hong Kong personalities working in areas likely to interest students in schools, e.g. fashion designers, songwriters. Besides obtaining questionnaire data on teacher reactions to materials, McNeill reports on some of the linguistic features that distinguish these spoken texts from materials written for language learners. Dismissing as irrelevant the view that authenticity resides solely in the language produced by native speakers, McNeill's discussion suggests that unscripted material of genuine interest in a local context offers a resource that is likely to prove suitable for developing listening skills, while stressing the need for more empirical research that directly involves "the most important personality in the whole equation: the learner".

The range of resources that can support language learning is much wider than is commonly appreciated in many environments, especially by students themselves. ***Stephen Ryan*** describes the initial views of language learners in a tertiary institution in Japan about the resources that might be available to them in future to maintain or advance their own learning of English. After warning that not all samples of "English" encountered in Japanese society would necessarily promote language learning, Ryan illustrates the many useful sources and activities that might be considered, and argues that awareness-raising in these respects has the important merit of extending the options available to learners in future situations where independent learning may become the only alternative to attrition.

A wider and richer view of "resources" emphasises the value of human resources rather than traditional concerns with "materials". The last two papers concentrate on human resources for teachers. ***Annie Mueller*** recounts the formation and development of an action research group at the University of Hong Kong, showing how this form of activity can not only demystify "research" but also tap into the potential enthusiasms, inventiveness and mutual inspiration of a group of language teachers. Mueller's account also documents the impact of this work on part of the curriculum (teacher support for self-access work as part of coursework), thereby confirming the symbiosis between personal and curricular development in such situations.

The demands of teaching in the school system can weigh heavily on teachers, so that the virtues of self-help groups may be hard to develop and sustain without external support. The paper by ***David Coniam, Sima Sengupta, Amy B.M. Tsui and Wu Kam-yin*** documents one form that this support might take, namely an interactive network that will enable teachers to ask for help on questions of grammar, share teaching ideas, contribute to and draw upon a bank of test items likely to be appropriate for their students. The paper introduces the TeleNex network which is currently being trialled in a number of Hong Kong schools, and which should become more widely available and appropriate as comments and feedback from teachers is incorporated into the various databases. The paper sets out the rationale, design and organisation of the network, which is intended to place technology at the service of better communication between language "experts" and schools. It is to be hoped that a future paper will evaluate the success of the first phase of this important initiative.

Desmond Allison

ENGLISH EXTENSIVE READING IN THE PRIMARY CURRICULUM - CURRENT PRACTICES AND NEW INITIATIVES

Vivienne Yu, Emily Chiu, Winnie Siu and Rachel Yau

1. Introduction

This paper describes a territory-wide survey of the use of children's literature, class readers and other extensive reading materials in Hong Kong primary schools. The survey was part of a larger, on-going project at the Institute of Language in Education (ILE) to develop an English extensive reading scheme for primary schools in Hong Kong, and was an attempt to find out whether extensive reading is currently practised in schools and teachers' to elicit opinions on the place of extensive reading in the primary curriculum.

In this paper, the findings of the survey are presented and the implications for the development of the Primary English Extensive Reading Scheme are discussed.

2. Background to the Survey

There has been a growing interest in English extensive reading programmes in Hong Kong schools in recent years. At the secondary level, there is a proliferation of reading schemes. Kwan (1988), in a survey of the use of extensive reading schemes in Secondary One, reported that of the 169 schools which responded to her questionnaire, 115 claimed they had some form of extensive reading programmes in their schools. Moreover, the ILE has also set up extensive reading schemes for public-sector schools in response to growing demand. The Scheme for junior secondary students is now well established, and at present there are 79 schools participating in the Scheme. In addition, the development of a reading scheme for upper primary pupils also began in the Institute in September 1993. Work on developing resource materials suitable for use by Primary 5 and 6 pupils is now under way.

As for secondary reading schemes, an important aim of the Primary Reading Scheme is to help pupils develop a reading habit. Feedback from teachers involved in the Secondary Reading Scheme and teachers on ILE Refresher Courses shows clearly that they believe the best chance for a reading habit to take root is to nurture it as early as possible. They point out that while it is good to have a reading scheme at Secondary One, it will be even better if there is a reading programme at primary level so that the habit can be formed earlier. As Southgate (1983:74) puts it when discussing ways of encouraging children to read, "the longer the establishment of the habit is delayed the less likely it is that it will ever be acquired."

Even though the idea of having English extensive reading at primary level has considerable support, hardly any information was available on whether there are extensive reading activities in primary schools at present. The survey was therefore undertaken to find out what the current practices are, and what primary teachers think about extensive reading. We also believe this background information is crucial if we are to develop a reading programme that can meet the needs of the schools. As our Scheme is designed for Primary 5 and 6 pupils the survey was targeted at upper primary level.

3. Aims of the Survey

The aims of the survey were as follows:

1. To investigate the extent to which extensive reading materials and extensive reading activity types are used at present at upper primary level. This would enable us to have a better understanding of the needs of schools as well as to build on what already exists when developing the Primary Reading Scheme.

2. To find out whether upper primary pupils have any difficulties in reading in English and if so, what these problems are. This would help us in the design of activities to increase pupils' reading proficiency.

3. To gauge primary teachers' opinions on the place of extensive reading in the curriculum. This would provide data on whether there is a positive climate for extensive reading at primary schools, and help establish the nature and extent of support needed.

4. Subjects and Methods

The data were collected by means of a questionnaire survey in October 1993. The questionnaire consisted of four main areas:

1. General information about the schools, including type (government, aided, private, co-educational, boys', girls' etc.), and number of classes in Primary 5 and Primary 6.

2. The use of class readers in Primary 5 and Primary 6. Schools were asked whether class readers were included in the curriculum, the type of activities used in teaching and their purpose in using the readers.

3. The use of class libraries at these two levels. Schools were asked whether there were class libraries in their schools, the kinds of materials included and the use these were put to.

4. English extensive reading in the primary curriculum. Schools were asked to express their opinions about extensive reading in the primary curriculum and to indicate whether they were interested in participating in the ILE Primary Extensive Reading Scheme. They were also asked to identify what problems (if any) their pupils experienced in reading.

The questionnaire was sent by post to 435 primary schools randomly chosen from the 852 primary schools in Hong Kong: The method employed was to select alternate schools listed in "Primary School Lists by District 1994-95" published by the Education Department. 294 questionnaires were returned, representing a response rate of 67.6%. These schools included a representative range of types (Table 1), and on average there were 3.5 classes at both Primary 5 and 6 levels.

The responses to the questions were computed for absolute frequencies and percentages. Full information appears in Tables 1-14 in the Appendix. The main findings are presented and discussed in the next section of the paper.

5. Findings and Discussion

5.1 The Use of Class Readers in Primary 5 and Primary 6

Of the 294 respondents, 210 (71.7%) prescribed class readers in their school textbook list for Primary 5 and 212 (72.6%) for Primary 6. Most schools used one to two class readers in a year (Tables 2 and 3). The three most frequently used activities in teaching class readers were questions and answers, reading aloud and silent reading. A variety of other activities were also used, ranging from role-play to dictation and tests (Table 4).

Schools were asked what aims they wanted to achieve through using class readers. Five aims were listed and the respondents were asked to rank them. The aim "To develop interest in reading" stood out significantly among the others. 170 respondents (85.5%) quoted this as either the most or second most important aim of using class readers in their schools (Table 5).

There is a dilemma here. As most schools use only one or two class readers a year, it would be difficult to help pupils to develop interest in reading and a reading habit. Other measures are obviously needed to address this issue.

5.2 The Use of Class Libraries in Primary 5 and Primary 6

Most respondents (280 schools or 95.9%) claimed that they had class libraries with English and Chinese books at both Primary 5 and Primary 6 (Table 6). The English materials and books in the class libraries included a range of types, with animal stories, fables, adventure stories, stories about daily life and fact books about hobbies and games being the most frequently cited examples (Table 7).

When asked how the books in the class libraries were used, schools indicated that most books were meant for pupils to borrow and read out of class, whether they be Chinese or English books. Although 60.6% of schools had library periods, it was unlikely that pupils would be given much help in their English reading since these periods were usually supervised by the school librarian or the class teacher rather than the English teacher. Only in 13 schools (4.7%) were pupils given time in the English periods to read English books, and even in these schools this was only done occasionally (Table 8).

The picture that emerges is that although primary pupils were not accustomed to reading English books on their own, little guidance and support were given. Moreover, our experience tells us that when pupils borrow books from the class library, they usually choose Chinese books. We may thus conclude that even though most schools have a selection of English books in their class libraries, this cannot adequately serve the purpose of promoting English extensive reading.

It is also interesting to find that although most schools allowed pupils to choose by themselves what titles to read from the class library, 73 schools (26.2%) sometimes or never let pupils choose their own reading materials or borrow books by themselves. In over half of these schools, books were given out randomly to pupils by the teacher (Tables 9a and 9b).

5.3 English Extensive Reading in the Primary Curriculum

The respondents were asked whether their pupils had problems with reading in English and if so, what they thought made reading difficult for their pupils. Seven frequently observed reading problems among pupils were listed and respondents were asked to indicate the degree of seriousness of each problem for their pupils on a four-point scale (from "most serious" to "not a problem for my pupils"). The overwhelming response was that teachers thought reading was difficult for the pupils, as shown in Table 10. The most serious problem encountered by the pupils, as reported by the teachers, was their inability to interpret implicit ideas in the text. Other difficult areas included tackling words, sustaining reading for a reasonable period of time, and to a lesser degree, following the main ideas of a text. The findings parallel the results of the IEA Reading Literacy study, which shows Hong Kong primary pupils had great difficulty with processing English texts, resulting in poor performance in reading tests (Cheung 1992).

The findings show clearly that more needs to be done to help primary pupils develop reading proficiency. Research studies on English extensive reading programmes at primary level seem to indicate that incorporating these programmes into the curriculum can help pupils to improve their reading ability. Elley's study of a nationwide extensive reading programme in the Fiji Islands shows that the 'book flood' classes made much greater improvement in all aspects of English than the control groups using an Oral English Syllabus based on audio-lingual approaches, and that the improvement was especially marked in reading (Elley and Mangubhai 1983). Similar positive language gains were also reported in two nationwide extensive reading programmes for primary pupils, one in Singapore (Ng 1988, Elley 1988) and the other in Brunei (Ng 1992). It will be interesting to see whether a reading programme for Hong Kong primary pupils will yield similar results.

The schools were also asked to rank seven given criteria for selecting books (Table 11). The three most highly ranked criteria were "interesting content for pupils", "appropriate length, structures and vocabulary items for pupils" and "good layout" (69.6%, 48.2% and 41.6% of the respondents respectively). A point of interest is that the criterion "language level matches exactly that of the coursebooks" was the fourth most highly ranked criterion on the list (29.4% of the respondents). This indicates that many teachers viewed extensive reading as a means to consolidate language taught in coursebooks rather than a channel to provide pupils with more exposure to the language.

When asked about their opinions on extensive reading, 277 schools (94.5%) believed that extensive reading was useful in helping their pupils to learn English (Table 12). They also believed that extensive reading could benefit their pupils in various ways (Table 13), including helping them gain confidence and interest in reading, develop reading skills as well as independent learning. Only 4 schools (1.6%) stated that they did not think their pupils could benefit from extensive reading.

Schools were also asked to indicate whether they were interested to join the Primary Reading Scheme that the ILE is developing. The result is very encouraging: 120 schools (40.8%) expressed an interest to join the Scheme and a further 20 (6.8%) would like more information (Table 14a). When asked how many English periods they thought should be allocated for extensive reading every week, 69.6% opted for one period a week, 13% suggested two periods, and another 13% half a period. The mean number of periods suggested was 1.03 (Table 14b).

However, it is worth noting that of the 277 schools which agreed that extensive reading was useful, only 120 expressed interest in joining the Scheme. Of course, it is possible that some schools are already running their own extensive reading programmes and do not see the need to join the ILE Scheme. But the more likely reason is that many teachers are still sceptical about allocating English periods to aid pupils' extensive reading. One can conclude that about 50% of the schools have seen the need for incorporating extensive reading into the curriculum, while extensive reading is still regarded by the other schools as an "optional extra", a "luxury" one can only afford when textbooks have been covered.

6. Conclusion - Implications for the Primary English Extensive Reading Scheme

The data from the questionnaire provide useful information for the development and design of the Primary Extensive Reading Scheme. Firstly, it is heartening to know that most teachers surveyed consider extensive reading as useful in helping their pupils learn English. They also seem to agree that the way this should be done is through helping pupils gain confidence and interest in reading and develop the ability to read independently. However, as discussed earlier in this paper, current practices in the majority of schools are unlikely to foster in the pupils a reading habit in English, since too little emphasis is given to extensive reading.

We hope the Primary Extensive Reading Scheme can help to redress the balance and serve as a breakthrough. Our plan is to fully incorporate the Scheme into the English curriculum and ask schools to allocate class time to it. This will ensure the pupils are constantly exposed to high-interest, comprehensible input and are also given help by their English teachers. The fact that 120 schools have indicated interest in participating in the Scheme shows clearly that even at this early stage, there is recognition that extensive reading should be accorded a much higher priority and become part of the English curriculum.

The survey also reveals that the teachers think their pupils have considerable problems with reading. These include, among others, reading the text word by word, not knowing how to guess words from context, not being able to understand meaning not explicitly stated, and finding the language too difficult. The implication is that besides providing pupils with high-interest books, we must also consider, in our design, ways to prepare our pupils to read books on their own. One way of doing this is to prepare "Help" cards that provide pre-reading activities to help them get into the book. These activities should aim at providing pupils with information of the general context and helping them use features such as pictures and the blurb to direct their understanding. Essential vocabulary items for understanding the book should also be introduced in the "Help" cards. In addition, "Activity" cards which pupils complete after reading a book will be useful in facilitating reading as this provides not only a reading record but also feedback to the pupils themselves on how well they have understood the book. Moreover, the "Activity" cards can stimulate pupils' imagination by asking them to give response to the characters, events, and ideas of the book. It is hoped that with the provision of high interest books as well as the "Help" and "Activity" cards, pupils will eventually be able to read books on their own and build up confidence and interest in reading extensively in English.

References

Cheung, Y.S. (1992). IEA (The International Association for the Evaluation of Educational Achievement) Reading Literacy Study: a preliminary report. *Hong Kong Reading Association Newsletter*, 21:6-8.

Elley, W. (1988). *The benefits of REAP (The Reading and English Acquisition Programme). An executive summary of the 1987 Evaluation* (Mimeo). Singapore: School Council Report SC/88/002.

Elley, W., & Mangubhai, F. (1983). The impact of reading on second language learning. *Reading Research Quarterly*, 19:53-67.

Kwan, K. (1988). *Project Report on Extensive Reading Schemes in Form/Middle One in Hong Kong 1986-87*. Hong Kong: Hong Kong Reading Association.

Ng, S.M. (1988). *The Formative Evaluation of REAP (The Reading and English Acquisition Programme)* (Mimeo). Singapore: School Council Report SC/88/003.

Ng, S.M. (1992). Improving English language learning for young speakers of other languages. Paper presented at the IRA (International Reading Association) World Congress on Reading, Maui, Hawaii.

Southgate, V. (1983). *Children who do read*. London: Macmillan.

Appendix

Table 1

General Information About Schools Responding to the Questionnaire

count	Boys'	Girls'	Co-ed	row total
Government		1	21	22 (7.5%)
Aided	2	12	245	259 (88.1%)
Private	1	3	9	13 (4.4%)
column total	3 (1%)	16 (5.4%)	275 (93.5%)	294 (100%)

Table 2

Class Readers Prescribed in the School Textbook List

count	Yes	No	total
P.5	210 (71.7%)	83 (28.3%)	293
P.6	212 (72.6%)	80 (27.4%)	292

Table 3

Number of Class Readers Prescribed

No. of	P. 5		P. 6	
books	count	percent	count	percent
1	120	58.3%	125	60.1%
2	79	38.3%	76	36.5%
>2	7	3.5%	6	3.0%

Table 4

Activities Used in Teaching Class Readers

Activity	Count	%
Questions & Answers	201	95.7
Reading Aloud	188	89.5
Silent Reading	182	86.7
Role-Play	129	61.4
Worksheets	67	31.9
Dictation & Tests	62	29.5

Table 5

The Aims of Using Class Readers

aim	count		cum. %
	most important	second most important	
To Develop Interest In Reading	143 (71.9%)	27 (13.6%)	85.5%
To Develop Reading Skills	35 (18.2%)	48 (25.0%)	43.2%
To Consolidate Grammar And Vocab. Learnt In Textbooks	13 (6.5%)	21 (10.6%)	17.1%
To Develop Skills In Reading And Writing	8 (4.0%)	9 (4.5%)	8.5%
To Develop Knowledge	5 (2.5%)	17 (8.5%)	11%

Table 6

Class Library For P.5 and P.6 Classes

	count	%
Yes	280	95.9%
No	12	4.1%

Table 7

Types of Materials Included in the Class Libraries

Books	Count	%
A. Fiction: Types		
Animal Stories	264	95.3
Fables, Legends, Fairy Tales	261	94.2
Adventures	233	84.1
Stories About Daily Life	223	80.5
Detective Stories	141	50.9
Thrillers and Ghost Stories	66	23.8
Love Stories	47	17.0
B. Non-fiction): Types		
Hobbies and Games	204	73.6
Facts of Life	171	61.4
Travels	127	45.8
Biographies	74	26.7
C. Others: Types		
Poems, Rhymes	159	57.4
Jokes, Riddles, Comics	113	40.8
Plays	33	11.9
Magazines and Newspaper	15	5.4

Table 8

How Books in the Class Libraries Are Used

	count Yes	No	total
To Be Read In English Periods	13 (4.7%)	266 (95.3%)	279 (100%)
To Be Read In Library Periods	169 (60.6%)	110 (39.4%)	279 (100%)
To Be Read Out Of Class	237 (84.9%)	42 (15.1%)	279 (100%)

Table 9(a)

Whether the Pupils Choose What To Read From the Class Library Themselves

	count	%
Yes	206	73.8%
Sometimes	68	24.4%
No	5	1.8%

Table 9(b)

If "Sometimes" Or "No", How the Books Are selected For Individual Pupils

	Yes	No	total
Teachers Give Out Titles Randomly To Pupils	47 (68.1%)	22 (31.9%)	69 100%
Teachers Select Different Titles For Pupils	29 (42%)	40 (58%)	69 100%

Table 10

To What Extent Each Of the Following Problems Makes Reading Difficult For the Pupils

1 = most serious;
2 = serious;
3 = not very serious;
4 = no problem for pupils

	1	2	3	4
Not Able To Understand Meaning Which Is Not Explicitly Stated In The Text	66 22.8%	156 53.1%	61 20.7%	3 1.0%
Do Not Know How To Guess Words From The Context, Including The Surrounding Text, Illustrations, Layout, etc.	66 22.8%	147 50.0%	73 24.8%	4 1.4%
Find The Language Too Difficult	64 21.8%	138 46.9%	81 27.6%	4 1.4%
Not Able To Sustain Reading For A Reasonable Amount Of Time	46 15.6%	148 50.3%	74 25.2%	11 3.7%
Always Read The Text Word By Word	42 14.3%	142 48.3%	93 31.6%	8 2.7%
Not Able To Follow The Main Ideas Of The Book	31 10.5%	117 39.8%	130 44.2%	11 3.7%
Always Read Back Unnecessarily	10 3.4%	102 34.7%	149 50.7%	21 7.1%

Table 11

Criteria For Selecting Books For Extensive Reading

criterion	count		cum. %
	most important	second most important	
Interesting Content For Pupils	101 (40.4%)	73 (29.2%)	69.6%
Appropriate Length, Structures & Vocab. Items For Pupils	59 (23.7%)	61 (24.5%)	48.2%
Good Layout	45 (18.0)	59 (23.6%)	41.6%
Language Level Matches Exactly That Of The Coursebooks	38 (15.3)	35 (14.1)	29.4%
Help Develop Pupils' Knowledge And Values	4 (1.6%)	12 (4.8%)	6.4%
Include Exercises And Activities	2 (0.8%)	5 (2.0%)	2.8%
Recommended By Publishers, Other Schools, Teachers, etc.	1 (0.4%)	3 (1.2%)	1.6%

Table 12

Whether The Respondents Think Extensive Reading Is Useful In Helping Their Pupils To Learn English

	count	%
Yes	277	94.5%
No	16	5.5%

Table 13

Perceived Benefits Of Extensive Reading By The Respondents

benefit	count		cum. %
	most significant	second most significant	
Gain Confidence & Interest In Reading In English	134 (53.8%)	47 (18.9%)	72.7%
Develop Reading Skills	44 (17.7%)	66 (26.5%)	44.2%
Develop More Independent Learning	34 (13.7%)	70 (28.1%)	41.8%
Improve Reading Speed	18 (7.2%)	32 (12.9%)	20.1%
Learn Useful Sentence Structures & Vocab. Items	16 (6.4%)	31 (12.4%)	18.9%

Table 14(a)

Whether The Respondents Would Like Their Pupils To Join The English Extensive Reading Scheme

	count	%
Yes	120	40.8
No	154	52.4
Not Decided	20	6.8

Table 14 (b)

If "Yes", Number Of English Period(s) Suggested Per Week For The Programme

No. of periods	count	%
1	80	69.6
2	15	13.0
1/2	15	13.0

INVESTIGATING LEXIS BEYOND THE MOST FREQUENT WORDS - PART 2

Norman Bird

1. Introduction

In the introduction to the AILA review on Vocabulary Acquisition Carter (1989:5) noted with the support of a sizeable bibliography that:

> The past decade has seen a considerable expansion of interest in vocabulary studies ... It can now be claimed that vocabulary is no longer a victim of discrimination by researchers who for a considerable period of time deemed syntax to be the sole core of processes of language development.

This interest in vocabulary acquisition has continued into the nineties and the one year since the 8th ILE Conference in 1992 has seen some notable advances. For example, in Hong Kong alone I.S.P. Nation gave a public lecture in City Polytechnic in November 1992 and his book *Teaching and Learning Vocabulary* became generally more available in the following year; 1993 also saw the publication of Pemberton, R. and Tsang, E.S.C. (eds.) *Studies in Lexis*, the working papers from the Second Seminar in Lexis held in the Language Centre, The Hong Kong University of Science and Technology (HKUST), and in June the third seminar on the same subject became a joint seminar on corpus linguistics and lexicology held in HKUST and the Guangzhou Institute of Foreign Languages.

After many years of comparative neglect, it seems that at last vocabulary has come to be recognized as one of the primary linguistic resources whereby meaning is encoded for the purpose of communication. In view of these recent developments, therefore, it is reasonable that reliable test instruments should be developed to measure the size of language learners' vocabularies as a step towards developing more efficient ways and means of increasing word power. It was with this in mind that the survey described in the first part of this paper was carried out, and it is hoped that the preliminary and necessarily limited findings described in this part of the paper will contribute to producing a still more efficient measuring instrument in the future.

The rationale, references, research methods and resultant problems for this research are described in full in Part One (Bird 1993). In brief, Part One describes an attempt to replicate the research described in Goulden, Nation and Read (1990) *How Large Can a Receptive Vocabulary Be?*, and produce a series of 50-word vocabulary tests of similar design in which each column of 10 words measures incrementally the mastery of 5,000 words in terms of frequency. The purpose of this research is:

1. to replicate and check the procedures previously used;

2. to use *The Oxford English Dictionary*, 2nd Edit. (1989) (OED2), as the resource corpus instead of *Webster's Third New International Dictionary* (1961) and *9,000 Words* (1983), since OED2 is both more up-to-date and reflects British rather than American English;

3. to produce materials that can be used in further research.

As a part of this research a questionnaire was produced consisting of the following:

1. a personal profile of the respondents, both native speakers (NS) and non-native speakers (NNS) regarding their mother tongue, age, formal qualifications and sex; for (NNS) of English an extra question concerning the number of years of exposure to English including time spent at school learning the subject is also asked;

2. one test from Goulden, Nation and Read (1990);

3. two tests attempting to replicate the Goulden, Nation and Read test, but based on OED2 (1989) and not Webster (1961 and 1983).

2. Responses to the questionnaire

Since the eighth ILE conference a total of 288 questionnaires were completed, returned and analyzed. The results in terms of the personal profile are presented in Table 1.

Table 1

Analysis of Questionnaire According to Personal Profile

a. English Native Speakers (NS) (total 78)

Quals.	Yrs. 20+	Yrs. 30+	Yrs. 40+	Yrs. 50+	-Total
Ph.D.	0	2	2	3	7
M.A.	0	9	19	11	39
B.A.	5	9	7	1	22
Others	2	4	0	4	10

b. Non-English Native Speakers NNS (Chinese) (total 201)

Quals.	Yrs. 20+	Yrs. 30+	Yrs. 40+	Yrs. 50+	Total
Ph.D.	0	0	0	0	0
M.A.	3	9	3	1	16
B.A.	9	43	22	1	75
Others	41	38	28	3	110

c. Non-English Native Speakers NNS (Non-Chinese) (total 9)

Quals.	Yrs. 20+	Yrs. 30+	Yrs. 40+	Yrs. 50+	Total
Ph.D.	0	0	1	1	2
M.A.	1	1	1	0	3
B.A.	0	0	1	0	1
Others	1	0	1	1	3

3. General Analysis of the Questionnaires

For the purposes of analysis, the rules of procedure described below were laid down after the first gross analysis of the returns.

3.1. Any particular category should contain at least 10 returns, otherwise the results should be ignored except as a description of general trends. As a consequence, the responses from NNS (non-Chinese) (Table 1c) are not included in the analysis for this research paper.

3.2. In order to limit the effect of variations in scores between tests the two following procedures were followed in the case of the analysis:

1. The two top scores out of 50 were added to give the general score out of 100 for the general analysis, and the third score was ignored.
2. As 87% of the NNS found the first 30 items of Test 1 to be more difficult than those in Tests 2 and 3, test results of the remaining 13% were ignored for the purposes of the detailed analysis.

3.3. As two discriminating factors in the personal profile (sex (cf. 4.4) and holding a Ph.D. vs. M.A.) were found to have no apparent effect on the test results, these two factors are ignored for the purpose of this small-scale research.

4. Detailed Analysis of the Questionnaires

The personal profiles were sorted and analyzed with respect to the following criteria: mother tongue, age, qualifications, sex and exposure to English (non-native speakers only).

4.1. **Mother tongue** was found to be the most crucial factor in ranking the results. The bar between NS and NNS results was 70%. One NNS speaker scored above this figure - namely, a 40+ year-old B.A. holder who had lived for over 20 years in the English-speaking world. Two NS scored below 70%; one obtained a score of 69% marginally below the 70% value; the other with 61% was a 30+ year-old Eurasian, neither of whose parents was a NS, who was brought up entirely in Hong Kong and, who had no formal qualifications beyond school certificate.

4.2. **Age** was found to be an important factor in the results both in the case of NS and NNS. Receptive vocabulary size remains stable or increases throughout that part of life measured in this research (cf. Table 2). The 2% variation between the scores of NS aged 30+ vs. 40+ years is assumed to be statistically insignificant.

Table 2

Comparative Analysis of Selected Tests

With the Single Variable of Age (and hence Experience)

a. Constant features: Chinese - B.A. - Female.

b. Constant features: Native-speakers of English - M.A.

Type	Age/Exp. 20+ (20)* %	Age/Exp. 30+ (25) %	Age/Exp. 40+ (25) %	Age/Exp. 50+ (25) %
a.	40	44	50	-
b.	-	82	80	90

*NB The figures in brackets (e.g. (20)) indicate the number of years of exposure to English.

4.3. **Formal qualifications** were also found to be an important factor in the scores of both NS and NNS, as can be seen from Tables 3 and 4.

289

Table 3

Comparative Analysis of Selected Tests With the Single Variable of Formal Qualifications

(Chinese Speakers of English)

Constant features: Chinese - Age (Experience) - Female.

Type	Age/Exp. 20+ (20)	Age/Exp. 30+ (25)	Age/Exp. 40+ (25)	Age/Exp. 50+ (25)
T.C.*	32	36	40	-
B.A.	40	44	50	-

*TC = Teacher's Certificate

In the case of NNS the most interesting feature to emerge from Table 3 is that the scores of 40+ year-old T.C. holders and 20+ year-old B.A. holders are identical. This suggests that holding the higher qualification of B.A. is equal to approximately 20 years of experience derived no doubt from the different quality of exposure to English that degree holders enjoy.

Table 4

Comparative Analysis of Selected Tests With the Single Variable of Formal Qualifications

(Native Speakers of English)

Constant features: Native-speaking English - Age.

Type	Age/Exp. 20+ (20) %	Age/Exp. 30+ (25) %	Age/Exp. 40+ (25) %	Age/Exp. 50+ %
B.A.	77	78	82	-
M.A.	-	82	80	90

Within the limits of this admittedly small sample composed solely of graduates, formal qualifications appear less important to native speakers than non-native speakers. In the case of NS holding a Ph.D. it was found that their scores did not differ from those holding an M.A., and consequently for the purposes of this research the factor of Ph.D. vs. M.A. was ignored.

4.4. Differences in **sex** were not found to affect scores. A small-scale comparison of the tests of two groups of 10 Chinese B.A. holders (aged 30-39) in which the only distinguishing feature was sex, showed a difference of only 0.2%; as this figure is of no statistical significance, the sex factor was ignored in attempting to form homogeneous groups for comparative purposes in this small-scale research.

5. Analysis of Tests

As mentioned in 3.2.2 the first general analysis of the returns revealed that 87% of the NNS found the first 30 items of Test 1 more difficult than those in Tests 2 and 3.

5.1. **A detailed analysis of the 3 tests (10 Chinese respondents)** is found in Table 5.

Table 5

Comparative Analysis of the results for Tests 1-3

From a Single Homogeneous Group of Chinese Respondents

Constant features: Chinese - Age: 40+ years - Exposure to English: 25+ years - Qualifications: B.A. - Profession: Teachers (ELT) - Sex: Both Males and Females.

Test	Q 1-10	Q 11-20	Q 21-30	Q 31-40	Q 41-50
1	95	52	18	16	09
2	100	82	48	21	13
3	98	79	41	13	07
Av. % Tests 2 & 3	99	80.5	44.5	18	10

Table 5 shows considerable discrepancies between the scores in Test 1 (Goulden et al.) and Tests 2 and 3 (Bird) especially in the first three columns (amounting to 30 items measuring the mastery of the 15,000 most frequent words).

Further investigation reveals that according to Thorndike and Lorge (1944) certain test words appear in inappropriate columns, as follows:

1. Column 1 (1-5,000 most frequent words): homage (6,000), colleague (7,000);
2. Column 2 (5,001-10,000 most frequent words): atrophy, broach, con, halloo, marquise, stationery, woodsman (beyond the 10,000 most frequent words).

5.2. **A detailed analysis of the 3 tests (10 NS respondents)** is found in Table 6.

Table 6 also shows considerable discrepancies between the scores in Test 1 (Goulden et al.) and Tests 2 and 3 (Bird). In the case of NS these discrepancies are not visible in the first two columns, as the test-words are known by virtually all the respondents. The scores for Test 1 in columns 3 and 4 are contrary to common sense, and suggest that this particular group knows the frequency group of words (15,000+) better than (10,000+).

Table 6

Comparative Analysis of the results for Tests 1-3

From a Single Homogeneous Group

of Native English-speaking (NS) Respondents

Constant features: NS - Age: 40+ years* - Qualifications: B.A. - Sex: M & F.

*3 respondents are 30+ but obtained the same median score as the 7 respondents in the 40+ group.

Test	Q 1-10	Q 11-20	Q 21-30	Q 31-40	Q 41-50
1	100	97	69	85	53
2	100	100	99	77	43
3	100	100	96	63	44
Av. % Tests 2 & 3	100	100	97.5	70	43.5

As Thorndike and Lorge (1944) is rather dated and based largely on school materials, the entries in Test 1 were also checked against the frequencies given in Hofland and Johansson (1982), the LOB list. It must be remembered, however, that LOB is based on a corpus of only one million running words of British English; furthermore the list has not yet been lemmatized, and thus measures the frequency of graphemes and not lexemes. LOB must, therefore, be used with extreme caution as an aid to obtaining an indicator of the order and trends in the frequency list of English words, but not as a statement of absolutes. A comparative analysis of the words in Test 1 and 2 (columns 2 and 3, i.e. lexeme frequency 5,001-15,000) is given in Table 7.

As mentioned in 3.2.2, in order to avoid problems arising out of variations in the difficulty of the three papers, and as 87% of the NNS found that the first 30 items of Test 1 to be more difficult than those in Tests 2 and 3, test results of the remaining 13% were ignored for the purposes of the detailed analysis.

This variation between tests affected NS less than NNS, i.e. only 68% found Tests 2 and 3 easier than Test 1, due in part to the fact that column 4 was easier than column 3. As a result of this discrepancy, it was decided to discontinue the use of Test 1 in future surveys, and for this research to use it with care, and only when necessary, if the size of an individual sample population to be studied was less than 10.

Table 7

Comparative Analysis of the Frequency According to LOB of the Words in Tests 1 and 2 (Columns 2 and 3)

Test and Column	Word and Frequency per Million
1 - 2	shrew, atrophy, con, halloo, marquise, woodsman (0), avalanche, firmament, broach (1) stationery (3).
2 - 2	swine, chink, ooze, filth (1), tributary (2), surf (4), idol (6), terminal (13), stationary (15), potential (42).
1 - 3	bastinado, countermarch, furbish, meerschaum, patroon, curricle, weta, bioenvironmental (0), regatta (4), asphyxiate (7).
2 - 3	misdemeanour, libertine, complicity, plethora (0), tentacle, dynamo, masticate/mastication, argent (1), whiff (2), disruption (3).

6. Conclusions

The 3 tests confirm what many may have long suspected but have not necessarily seen formally measured, namely that the most important factor in receptive vocabulary is being a NS. Age and hence exposure to English are also clearly important, but the 'advantages of age' are offset by the possession of formal educational qualifications and all that this implies; this factor is of greater importance to NNS (Chinese) than it is to NS. Difference in sex appears to play no role as a determining factor, and some may choose to discount it in similar research in the future.

7. Implications

Implications from the results of these tests should only be drawn with extreme caution. There is evidence that the receptive vocabularies of NS are general larger than those of NNS, but possessing a large receptive vocabulary is probably nothing more than an indicator of extended exposure to the language especially in the early years of life when the mind is at its most receptive. This advantage that NS have over non-native speakers can be compensated for in part by age, formal qualifications and extended periods of exposure to the language, and it is encouraging to observe that education, especially in the form of studying for university degrees, demonstrably builds up the receptive lexical resources that non-NS have at their disposal; the extent to which this is possible, however, can only be deduced with difficulty from the results of this piece of small-scale research comprising less than 300 scores, and the tests must clearly be made more sophisticated and given under carefully controlled conditions to larger samples of the population of NNS.

The test also raised two particularly interesting questions, namely:

1. What guarantee does the tester have that the respondent is answering the questions honestly?

2. What is meant by 'knowing a word'?

The answer to the first question is that the tester can never be absolutely certain that tests are being answered honestly, but this does not necessarily invalidate all them. Firstly, respondents gain no benefit from answering the questions dishonestly, as they are informed that the test is carried out for statistical purposes only, and if respondents so wish, they can remain completely anonymous. Furthermore, as certain patterns begin to emerge in the answers, as more and more papers are tallied, unlikely answers rapidly become apparent, and consequently, this potential problem soon ceases to be a real problem at all.

The second question as defining the concept of 'knowing a word' is discussed at length in Nation (1990:30-32); from his Table 3.1 (p.31) it is clear that, if we consider the 18 features included listed there, only two features are addressed by the tests considered in this paper, namely:

1. Written form R* What does the word look like?
2. Meaning R What does the word mean?

R* = (receptive vocabulary), i.e. not productive vocabulary.

One final point of interest may be worthy of mention here, although it does not appear in the test results. It was observed that among those tested who scored approximately 65%, i.e. many NS and a few NNS, the question was often raised as to whether a word could be regarded as 'known' if, although it had never been encountered before, it was felt the meaning of the word would create no problem, if it were encountered in a proper reading context, e.g. aquose - adjective, technical, scientific, probably something to do with water, e.g. full of water. To such a question the answer was given that the word is known and awarded and one mark was scored. Two points of interest arise here. Firstly, when does this degree of lexical self-confidence and sophistication begin to appear, because there can be little doubt that once it does a learner's vocabulary increases very rapidly bringing with it obvious results in such skills as reading. Secondly, is it possible to bring language learners to this point more rapidly than at present by pedagogic means, i.e. by making students increasingly 'root conscious'.

The tests presented here and the results that they have produced must, therefore, be considered within the wider context of vocabulary testing in general where they can, in fact, play a useful although limited role as a part of a larger battery of tests.

References

Bird, N. (1993). Investigating lexis beyond the most frequent words - Part 1 In N. Bird, J. Harris and M. Ingham (eds.), *Language and Content* (pp. 437-450). Hong Kong: Education Department, Government Printer.

Bird, N. (1990). *A first handbook of the roots of English.* Jersey C.I.: Lapine Education and Languages Services Ltd.

Goulden, R., Nation P. & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics,* 11(4):341-357.

Hofland, K. & Johansson, S. (1982). *Word-frequencies in British and American English.* Bergen: The Norwegian Computing Centre for the Humanities (NAVF).

Nation, I.S.P. (1990). *Teaching and learning vocabulary*. Boston, MA: Heinle and Heinle.

Oxford English Dictionary, (1989). (2nd edit.) Oxford: Oxford University Press.

Pemberton R., & Tsang, E.S.C. (1993). *Studies in lexis*. Hong Kong: The Hong Kong University of Science and Technology Language Centre.

Thorndike, E.L. & Lorge, I. (1944). *The teachers's word book of 30,000 words*. New York: Teacher's College, Columbia University.

Webster's Third New International Dictionary, (1961). Springfield, MA: Merriam-Webster Inc.

9,000 Words, (1983). Springfield, MA: Merriam-Webster Inc.

Appendix

Questionnaire

I would be pleased if you would kindly complete the following questionnaire and the three attached vocabulary tests.

All the information will be kept strictly confidential and used for statistical purposes only.

Please circle the appropriate word.

1. Sex: Male Female

2. Age: 20+ 30+ 40+ 50+

3. Mother tongue: English Chinese Other

4. If "Other" in "3" above, please state your native language
   ....

5. Qualifications: Doctor, e.g. Ph.D.
   Master, e.g. M.A.
   Bachelor, e.g. B.A.
   Secondary School Certificate
   Other (please state) ....

6. If you are not a native speaker of English, please state the number of years you have used English, including your periods of studying the language.

   1+ 5+ 10+ 15+ 20+ 25+ yrs

Tick the words you know. Add the number of ticks and give the total at the end of each test.

Test 1*

1 bag
2 face
3 entire
4 approve
5 tap
6 jersey
7 cavalry
8 mortgage
9 homage
10 colleague
11 avalanche
12 firmament
13 shrew
14 atrophy
15 broach
16 con
17 halloo
18 marquise
19 stationery
20 woodsman
21 bastinado
22 countermarch
23 furbish
24 meerschaum
25 patroon
26 regatta
27 asphyxiate
28 curricle
29 weta
30 bioenvironmental

31 detente
32 draconic
33 glaucoma
34 morph
35 permutate
36 thingamabob
37 piss
38 brazenfaced
39 loquat
40 anthelmintic

41 gamp
42 paraprotein
43 heterophyllous
44 squirearch
45 resorb
46 goldenhair
47 axbreaker
48 masonite
49 hematoid
50 polybrid

Test 2

1 ball
2 dead
3 loose
4 royal
5 stomach
6 veil
7 screw
8 fee
9 mask
10 beak

11 idol
12 tributary
13 potential
14 swine
15 chink
16 stationary
17 ooze
18 terminal
19 filth
20 surf

21 tentacle
22 dynamo
23 whiff
24 disruption
25 misdemeanour
26 libertine
27 mastication
28 complicity
29 plethora
30 argent

31 copulate
32 paradigm
33 cadge
34 aquaplane
35 antiphon
36 acrostical
37 shimmy
38 pomander
39 basqueless
40 parameter

41 slurry
42 ska
43 prunella
44 glycerose
45 pessimum
46 remanence
47 rhinocerotid
48 secant
49 minikin
50 tansy

Test 3

1 day
2 fear
3 memory
4 plane
5 female
6 twist
7 breeze
8 pluck
9 jam
10 pulse
11 bluff
12 hesitation
13 exit
14 treatise
15 feverishly
16 sill
17 giggle
18 swerve
19 clod
20 innovator
21 patriarch
22 tumour
23 barb
24 whetstone
25 chamois
26 sty
27 hydraulics
28 addle
29 tactile
30 flageolet

31 parse
32 spew
33 reginal
34 aureate
35 corral
36 debrief
37 sinistral
38 ablation
39 carriole
40 parhelion
41 possum
42 scarification
43 muscatel
44 aquose
45 erythraemia
46 pelagian
47 irredentist
48 doweral
49 helvellic
50 farinulent

*Goulden, R., Nation P., and Read, J. (1990). *How large can a receptive vocabulary be?* (Test 2) Applied Linguistics, 11(4):359.

THE FUTURE ROLE OF EFL TEXTBOOK RESOURCES IN HONG KONG

Bob Adamson and John C.K. Lee

1. Introduction

The teaching and learning of English is about to undergo radical change in Hong Kong. The forthcoming introduction of the Target-Oriented Curriculum (TOC) calls for new approaches to planning and organising learning experiences, and to assessment procedures.

As with all instances of curriculum renewal, the resources that are made available will go a long way to determining the eventual success or failure of TOC. Therefore, it is pertinent to examine a traditional mainstay of language teaching resources, the textbook. What exactly is a textbook? What purposes does it serve? Is there a place in TOC for the textbook? What, if anything, should or can take its place?

In this paper we shall briefly outline the evolution of English language textbooks in Hong Kong within different educational values systems. This historical perspective will inform our discussion concerning the production of suitable resources to support the TOC initiative. We shall try to identify a practical role for textbooks that is compatible with the socio-political, educational and economic forces that shape the design, implementation and enactment of TOC.

2. Theoretical framework

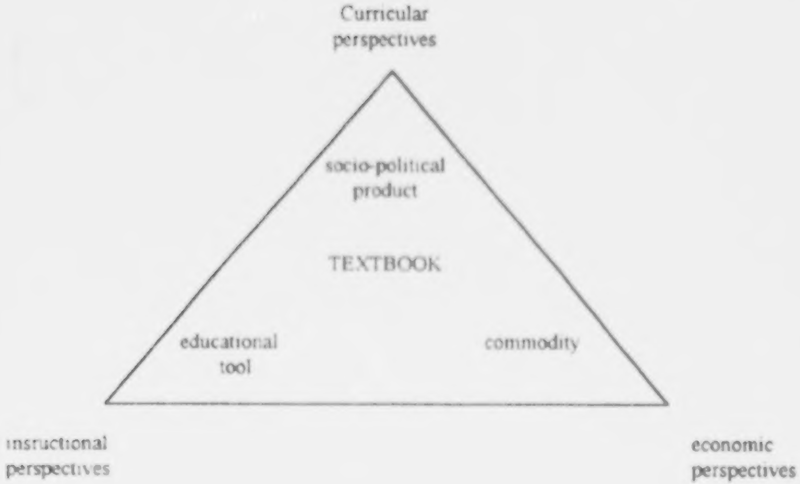
For our investigation of the role of EFL[1] textbooks in Hong Kong, we will adopt a triangular theoretical framework (figure 1). This framework, to which a third, historical, dimension will be added, enables us to examine textbooks as socio-political, educational and economic products.

The three perspectives, socio-political, educational and economic, represent the interests of, respectively, those who plan the curriculum (usually government agencies), those who implement the curriculum (teachers and students), and those who resource the curriculum (principally commercial publishers).

The curriculum in general is laden with socio-political values (Clark 1987), and the English language syllabus in particular is shaped by views about language learning. These values and views will be reflected, to varying degrees of fidelity, by the nature and use of textbooks. In time, values systems have altered and views about language learning have been reassessed. One would expect to find that the nature and use of textbooks have likewise undergone change.

Figure 1

Theoretical framework



'Instruction', in this paper, carries the broad meaning of 'strategies intended to facilitate learning'. The features of textbooks as instructional tools are defined by their contents and the ways they are actually used by teachers and students.

Considering the textbook as a commodity brings economic factors into play. Commercially produced textbooks are subject to market forces which include the profit motive, the laws of supply and demand, the availability of resources (including human resources), and the available technology.

We shall apply the theoretical framework to three periods of English language teaching in Hong Kong. Borrowing from Skilbeck's (1982) labels for Western values systems, we shall label the Hong Kong periods as Classical Humanism, Reconstructionism and Progressivism, as Hong Kong has tended to import its educational reforms from the United Kingdom. The periods will be crudely demarcated as 1878-1952, 1952-1995 and post-1995 respectively. We say 'crudely' because these dates mark significant educational decisions or the commencement of curriculum renewal: there often follows a time-lag before widespread acceptance and implementation of change occurs.

3. Classical Humanism: 1878-1952

In 1878, a Conference on the Teaching of English at Government Schools concluded that the teaching of English should be accorded the highest priority in education (Ng 1984:70). Education in schools that taught English was generally reserved for the elite and wealthy (ibid:69). English language teaching focused on the acquisition of rules about language, accessed through linguistic analysis of literary extracts. The syllabus sequenced language forms from what was perceived to be simple to what was perceived to be difficult. This approach to language education is consistent with the general properties of Classical Humanism which Clark (1987:3) summarises as:

> ... elitist, concerned with the generalizable intellectual capacities and with the transmission of knowledge, culture and standards from one generation to another.

Instruction consisted of the technical implementation of the syllabus, whereby the teacher, as possessor of the appropriate knowledge, divulged it to the students, who used the techniques of memorisation, translation, analysis, classification and reconstruction of given rules and principles to acquire and apply the knowledge. This acquisition was then assessed according to norm-referenced criteria.

Textbooks were employed primarily as source books of culture and language analysis. Typical kinds of textbook would be the reader, incorporating a collection of literary extracts complemented by *explications de texte*, and a grammar book, with rules and exercises. A supplementary writing book, for dictations, calligraphy and elementary composition was also used. The subject matter of lessons, usually taught by teachers with little or no training (Sweeting 1993), would be dictated by the contents of the textbook.

As commodities, these textbooks were usually imported from publishing houses in Britain, such as Macmillan and Company, T. Nelson and Sons and Blackie & Son (Fong 1975:68-74), or occasionally published in Hong Kong through subsidies from missionary organisations or educational authorities. Textbook writers were native speakers, most having no connection to Hong Kong. The size of the market for EFL textbooks, numbering around ten thousand towards the end of the nineteenth century (Ng 1984:163), rendered it an unprofitable proposition for publishing houses.

4. Reconstructionism: 1952-1995

The characteristics of the second values system, Reconstructionism, with its view of education as a dynamic means for egalitarian social reform, mark it as a clear departure from the elitist, conservative tendencies of Classical Humanism. This shift accorded well with the imperative need to develop a workforce for Hong Kong's growing manufacturing industry in the post-war period. At the same time, the population was swelling rapidly, necessitating an expansion of basic education.

A series of reports addressed the issue of language teaching. The Burney Report (1935) emphasised the need to develop vernacular education at the expense of English, which would be taught for vocational needs. Burney's recommendations could not be implemented until after the Second World War, when a serious reappraisal of education in Hong Kong was made. In 1952, a report on English teaching in schools by a Committee on Higher Education found serious weaknesses and, as a consequence, reforms were undertaken (Bickley 1987). The Llewellyn Report (1982) stressed the importance of 'genuine bilingualism' (Llewellyn et al. 1982:29) and proposed various measures to improve the quantity and quality of education in Hong Kong.

Initially, the form of Reconstructionism was very mild. A move towards egalitarianism was manifested in the improved provision of education, but serious efforts to use education as a means to effect fundamental social change only began after the signing, in 1984, of the Draft Agreement between the People's Republic of China and the United Kingdom concerning the future of Hong Kong (Morris 1992).

The mildness of Hong Kong Reconstructionism, particularly in the post-war period, is especially evident in the contents of EFL textbooks. The Classical Humanist literary texts, with their cultural function, were replaced by dialogues, stories (often simply anecdotes) and passages of general interest. Little attempt was made at consciousness-raising or promoting egalitarianism, even after the watershed of 1984. Present-day textbooks devote occasional pages to social issues such as the environment or the problems of the handicapped, but the principal function of Reconstructionist EFL textbooks has been to promote language learning as an end, rather than as a means to other educational goals.

As with Classical Humanism, the English syllabus is subject-centred, but defined through a synthesis of theory and practice, and implemented through critical enactment, whereby teachers adopted a more active role in the selection of subject matter. In terms of instruction, the rule-based, direct transmission tenets of grammar-translation were replaced by a more pragmatic view of language learning which stressed language use rather than knowledge about language. The goal was for students to achieve communicative competence in oral and written language.

A number of instructional strategies were used, including the Direct Method, Audio-lingualism and Oral-structuralism (Bickley 1987). The scope of EFL textbooks widened to incorporate communicative language serving as a model or a stimulus, with varying degrees of authenticity (Ng 1993), for communicative output by learners.

The textbook, as a static, writing-based, two-dimensional medium has severe limitations for the purpose of focusing on communicative language in its diverse oral, aural, and visual manifestations. These limitations were eased by technological innovations, resulting in the development of a wide range of aural and audio-visual resources. One might have expected that the place of the textbook as the central resource in language learning be threatened by such technological advances. However, a recent study (Richards, Tung and Ng 1992) shows that teachers in Hong Kong rely heavily on textbooks as their principal instructional tool.

We would suggest three principal reasons for the dominance of textbooks over other forms of resources. The first is the weight of tradition: the book has historically been the central learning resource. The second reason is the permanence and convenience offered by the book as a printed, centralised and portable collection of materials. Thirdly, as a commodity, textbooks command a larger market and are less prone to illegal copying than resources such as cassette and video tapes.

With the growth in population, the market for textbooks increased considerably during the Reconstructionist period, resulting in the establishment of local and international publishing houses in Hong Kong. The lack of experienced, professional textbook writers has been overcome by commissioning overseas textbook writers, language specialists employed by the Education Department and academics in local tertiary institutions to produce materials. More recently, measures to strengthen language education and teacher education have enabled publishers to engage expatriate and - a notable departure - local Chinese teachers in Hong Kong as EFL textbook writers.

Quality control of textbooks is carried out by a committee under the auspices of the Education Department, which requires publishers to submit all textbooks for approval.

5. Progressivism: post-1995

Throughout the eighties, much of Hong Kong's manufacturing base emigrated to Guangdong and other parts of south-east Asia. As a result, Hong Kong developed information-based and service industries to fill the void. At the same time, rapid technological advances, most notably computerisation, and the globalization of trade, including tourism, created a demand for a highly skilled and linguistically competent workforce.

Recent political movements experienced particularly in Western countries, such as democratisation, human rights issues, and equality for all sectors of a population, have also been echoed in Hong Kong. These trends in Hong Kong have grown from its political, economic and communication links to the West, improved educational and living standards, and sensitivity in relation to its hand-over to the People's Republic of China in 1997.

As a consequence of these global changes, and of the development of new theories of learning, Hong Kong plans to adopt a Progressivist approach to the curriculum in 1995, in the form of the Target-Oriented Curriculum (formerly known as Targets and Target-Related Assessment).

Progressivism focuses on the needs of an individual as an intellectual, emotional and social being. Unlike Classical Humanism or Reconstructionism, it is essentially learner- and learning-centred. It does not envisage the 'top-down' transmission of a value-laden, subject-centred curriculum as its chief *raison d'être*. Instead,

> [k]nowledge is not seen as a set of fixed facts, but as a creative problem-solving capacity that depends upon an ability to retrieve appropriate schemata from a mental store, to utilize whatever can be automatically brought to bear upon a situation, and to bend existing conceptual structures to the creation of novel concepts that offer a working solution to the particular problem in hand. (Clark 1987:50-51)

Progressivism is realised in language teaching and learning through task-based programmes such as those pioneered by Prabhu in the Bangalore Project (Prabhu and Carroll 1980). As with Reconstructionist approaches, it stresses the communicative nature of language learning. However, there are two important differences.

Firstly, Progressivism does not view language learning as an end in itself. Rather, it holds that language is learnt *for*, and *through*, personal growth. It follows that the English syllabus does not consist merely of language items or concepts graded according to the perception of whether they are 'easy' or 'difficult'. Instead, it aligns a unified spiralling progression of cognitive goals and targets relating to the language contents, processes and products. The individual tasks are thematically linked in units, which, in turn, form part of a unified module. Assessment takes the form of criterion-referencing against holistic communicative targets.

Secondly, it is the TOC tasks, chosen to match the needs and interests of the specific group of learners, that determine the lexical, grammatical and other linguistic elements that learners will need, whereas formerly it was the language syllabus that determined the exponential communicative tasks.

Progressivist instruction follows the mutual adaptation model, which means that teachers and learners take most of the responsibility for the choice, design and conduct of learning experiences. Many of the experiences involve pair work, group work and individual work on projects and tasks, although supporting language-focused exercises and practice allow for some teacher-centred teaching. Another feature of instruction is learner-independence in selecting their own paths of learning and use of English.

The learner-centred approach, in which instruction caters to individual needs and interests, has important implications for the future of EFL textbooks. To date, textbooks, whether Classical Humanist or Reconstructionist, have carried a graded language syllabus, tailored, in many cases, for particular groups of students (such as those in Band 1 schools, for instance). The learner-centred TOC syllabus, on the other hand, envisages courses with sophisticated cognitive and linguistic grading criteria, designed for much smaller groups, or even individuals. Clearly, then, subject-centred textbooks designed for a general market are incompatible with learning-oriented methods that cater to the individual needs of learners.

Richards (1993), who was not referring specifically to the Hong Kong situation, identifies three other ways in which the use of commercially produced textbooks might have a negative impact, namely:

- the lack of local content
- the reification of textbooks
- the deskilling of teachers

The lack of local content may not, at first glance, be considered an important issue for Hong Kong, which occupies such a relatively small geographical area. However, within the territory, there is actually considerable diversity in terms of locality and social background of learners; this diversity is seldom addressed in textbooks produced 'for the Hong Kong market.'

Reification of textbooks, whereby teachers ascribe a superior status to them, is inappropriate in a learner-centred culture. TOC emphasises the central role of teachers as course designers in meeting the specific needs and interests of learners. It follows, therefore, that if teachers abdicate their decision-making role in favour of relying on the textbook, they are deskilled and play a lesser part in the instructional process.

Theoretically, therefore, it would appear that textbooks are too clumsy an instructional tool for a Progressivist values system. Ideally, they would be replaced by teacher- or even learner-produced modular resources that match the needs and interests of individual learners, or, at most, small groups of learners. Indeed, there has recently been a rapid growth internationally in resource books for teachers, comprising 'recipes' or suggested activities on which teachers can draw when constructing schemes of work.

Therefore, if TOC to be implemented according to its theoretical ideals, there is a need for teachers to act as materials designers and producers. What is not clear is the current circumstances of teachers in Hong Kong permit them to fulfil this function.

Although much excellent work is done in some schools in materials preparation, the study by Richards, Tung and Ng (1992) found that less than one third of the teachers surveyed made significant use of self-prepared materials. Several reasons for this may be proposed.

Firstly, the Chinese Neo-Confucian tradition favours a text-based approach to language study. This tradition could influence not only teachers, but also parents and students. Secondly, Education Commission Report Number 4 reveals that 46% of primary and 44% of secondary English teachers are not subject-trained (Education Commission 1990), which would explain a predisposition amongst such teachers towards textbook-dependence. (Cf. 1991 figures reported by Tsui in Coniam et al. in this volume: editor's note.) Thirdly, large classes of around forty students and a heavy teaching load reduce the available time for materials preparation.

Furthermore, there are potential difficulties with 'in-house' teacher-produced resources. Quality control presents a major problem. Not all teachers can be expected to have the necessary creativity, linguistic competence and instructional design skills: nor is there any guarantee that teacher-produced resources will automatically be more successful than commercial resources.

Schools do not have the printing or reproduction facilities to produce attractive materials of a professional standard, nor the capacity to cope with the large amount of production work that would arise if all departments were attempting to create their own materials. Another drawback would be that the wheel would be re-invented on numerous occasions unless an effective mechanism existed to allow schools to share their materials.

One solution might be to encourage publishers to produce a wide range of individual modules in booklet form, from which teachers could select or adapt suitable materials to form a coherent course for their particular circumstances. However, to produce these booklets with textbook-standard artwork, paper and editorial input would incur a higher unit cost than the larger-scale textbook. Booklets would also be more vulnerable to illegal photocopying, thus reducing sales and increasing prices.

6. The Future Role of Textbook Resources

We believe that published textbook resources offer a viable solution to the problem of resourcing the TOC. The term 'textbook resources' represents the nature of EFL instructional materials since Reconstructionist times. It refers not only to the textbook that forms the core of the materials, but also to the complementary resources, such as workbooks, audio cassette and video tapes, that are necessary because of the complex nature of language learning.

There are three principal reasons why published textbook resources still have a role to play in a Progressivist values system: durability, flexibility and economy.

To date, the textbook as core instructional material has proved to have qualities of endurance in resisting challenges from other forms of resources, and to be adaptable in times of educational change. The nature of the post-war EFL textbook made the transition from providing a collection of literary texts to serving as an eclectic language resource, despite being ill-fitted to convey the oral medium of communication. The economic advantages of the textbook, which contribute in large part to its enduring quality, are reinforced by the reified status accorded to textbooks by teachers, which, although educationally undesirable, promotes confidence among learners, teachers and parents in its use.

To allow flexibility in the use of textbook resources, two important criteria need to be met. Firstly, the textbook resources must be carefully designed and written to reflect the basic tenets of TOC. Secondly, teacher education should seek to enskill teachers to handle textbooks in ways appropriate to individual learning situations.

Progressivist textbook resources would consist of several modules. The modules should be free-standing, but, taken together, would form a coherent series to cover an individual keystage of TOC. The approach, following TOC principles would be thematic, in that the tasks would dictate the language syllabus and not, as previously, *vice versa*. The modules would be graded in terms of cognitive and linguistic demands upon learners. However, there would be scope for adaptation by teachers, so that modules would not necessarily have to be followed in a predetermined order.

The materials would comprise a core textbook, supported by complementary resources. With the socio-political emphasis on whole-person, experiential learning, the materials in the core textbook would include experiential text, intended to promote whole-person response, as well as offering the wherewithal to create holistic learning experiences in the form of tasks.

The complementary resources would have two functions. The first would be to create alternative experiences through a different medium. This would involve the provision of video and audio cassette tapes, or whatever hardware and software may arise from technological developments. The second function would be to cater for individual differences in learners, by providing remedial, supporting and extension work through language and skills-centred exercises.

Experiential texts, in a range of oral, visual and written genres, would pursue the Progressivist principle of language learning *for* and *through* personal growth by engaging the learner in critical reflection on cross-curricular political, social, intellectual and emotional issues: EFL textbook resources would no longer be permitted to adopt the isolationist stance of language learning for its own sake. Textbook resources hold an advantage over teacher-produced resources, in that researching suitable cross-curricular materials and producing high-quality resources are very time-consuming and expensive operations. Comprehensive support for teachers should also be provided, so that the textbook and other resources also perform a teacher education function. Once again, the form of teacher education provided by the textbook resources must be consistent with the tenets of Progressivism: teachers should be encouraged to become reflective practitioners. Notes for teachers could include sample teaching plans and advice on methodology, as well as background information about cultural and linguistic matters likely to arise. They should also present ways of adapting the materials to meet individual needs.

Formal pre-service and in-service teacher education must prepare teachers to be careful selectors and adaptors of published materials. The choice of textbook resources is a crucial decision, requiring informed opinions. Teachers can be introduced to selection criteria, such as the *Catalyst* checklist (Grant 1987).

Teacher education programmes can equip teachers to use the textbook resources in three different ways. Firstly, teachers could adopt the textbook resources for class use uncritically. Although far from ideal, such use would ensure that the learners have access to good quality and well-structured resources, whilst providing support and education for teachers who are, initially at least, inexperienced, unqualified or unable for other reasons to play a more active role in materials design.

At the other end of the spectrum, teachers could be enskilled to manage without textbook resources and be responsible for their own materials production. In these circumstances, the textbook resources are not used in the classroom, but made available in a resource centre. The modules would then serve as exemplars for the teachers.

The third use would see the textbook resources adopted for classwork, but the contents would be adapted, omitted or replaced as necessary by the teacher to match the needs and interests of the learners. This solution would allow teachers to use their professional skills, whilst, at the same time, releasing them from the total burden of materials production.

In economic terms, publishers would relish the prospect of a large potential student market[2] that the continued use of textbook resources would permit. One benefit would be high standards of production that the lower unit costs of textbooks allow. To ensure that textbook resources meet the specific needs of the Hong Kong market (or, preferably, individual sectors of that market) publishers should continue recent trends by encouraging participation by local teachers in all stages of planning, writing and producing textbook resources.[3]

At the same time, quality control is an essential factor if textbook resources are to fulfil their role adequately. These resources must be worthy of the confidence placed in them by learners, parents and teachers. Publishers need to ensure that high standards of design and instructional content are maintained. One way to achieve this would be to have drafts of the materials reviewed by consultants, including curriculum specialists and experienced materials writers, as well as teachers and learners. The Education Department's role in controlling the quality of textbooks should also continue and be strengthened, thus making sure that textbooks adhere to the principles of language learning according to the philosophy and framework of TOC.

7. Conclusion

The planned introduction of the learner- and learning-centred Target-Oriented Curriculum in Hong Kong in 1995 threatens generalist, subject-based textbook resources with extinction. This paper suggests that a role still exists for EFL textbook resources within TOC. Such resources would typically be:

- task-based, organised thematically
- graded according to cognitive and linguistic principles
- experiential, using texts that promote whole-person growth
- teacher-educational, promoting reflective practices
- produced with the involvement of local teachers

The continued use of textbook resources would provide incentives to two key stakeholders in curriculum renewal: teachers and publishers. For teachers who view the role of materials producer with trepidation, published textbook resources offer solace, by entrusting resource production to those with a creative streak and the time and energy to commit to such projects. Publishers, with continued access to a large market, would be empowered to produce well-made commodities which, thanks to appropriate quality control procedures, would constitute high-quality instructional tools.

As such, Progressivist textbook resources represents a compromise between theoretical socio-political idealism and practical educational and economic realism.

Notes

1 EFL = English as a foreign language. A discussion whether English is learnt in Hong Kong as a *foreign* or a *second* language is beyond the scope of this paper.

2 Currently there are approximately 100,000 primary students and 80,000 secondary students in Hong Kong.

References

Bickley, V. (1987). Developments in English Language Teacher Education. In R. Lord & H.N.L. Cheng (eds.), *Language Education in Hong Kong* (pp. 187-218). Hong Kong: The Chinese University Press.

Clark, J.L. (1987). *Curriculum renewal in school foreign language learning*. Oxford: Oxford University Press.

Education Commission. (1990). *Education Commission Report Number 4*. Hong Kong: Government Printer.

Fong, M.Y. (1975). *The first hundred years of Hong Kong education*. Hong Kong: China Learning Institute (in Chinese). Translation of pp. 68-74 in Sweeting (1990) (pp. 265-269).

Grant, N. (1987). *Making the most of your textbook*. Harlow: Longman.

Llewellyn, J. et al. (Hong Kong Government). (1982). *A perspective on education in Hong Kong, report by a visiting panel*. Hong Kong: Government Printer.

Morris, P. (1992). Preparing pupils as citizens of the Special Administrative Region of Hong Kong: An analysis of curriculum change and control during the transition period. In G.A. Postiglione (ed.), *Education and society in Hong Kong: toward one country and two systems* (pp. 117-145). Hong Kong: Hong Kong University Press.

Ng, L.N.H. (1984). *Interaction of East and West: development of public education in early Hong Kong*. Hong Kong: The Chinese University Press.

Ng, P.E. (1993). How communicative are English language textbooks in Hong Kong? Unpublished B.Ed dissertation, University of Nottingham.

Prabhu, N.S. and D.J. Carroll (1980) Syllabus design and methodology. *Newsletter (Special Series)*, 1(4):1-20. Bangalore: Regional Institute of English.

Richards, J.C., Tung, P. & Ng, P. (1992). The culture of the English language teacher: a Hong Kong example. *RELC Journal*, 23(1):81-102.

Richards, J.C. (1993). Beyond the text book: the role of commercial materials in language teaching. *RELC Journal*, 24(1):1-14.

Skilbeck, M. (1982). Three educational ideologies. In T. Horton & P. Raggatt (eds.), *Challenge and change in the curriculum* (pp. 7-18). Sevenoaks: Hodder and Stoughton.

Sweeting, A. (1990). *Education in Hong Kong Pre-1841 to 1941: fact and opinion*. Hong Kong: Hong Kong University Press.

Sweeting, A. (1993). Historical perspective: the long march towards teacher professionalism in Hong Kong. In D. Chincotta (ed.), *Professionalization and education* (pp. 1-40). Education Papers No.15. Hong Kong: Faculty of Education, The University of Hong Kong.

CONCORDANCING FOR SCHOOLS: PROBLEMS AND POTENTIAL

Valerie Pickard, Kenneth Chan and Janice Tibbetts

Introduction

Since the publication of the Collins Cobuild English Language Dictionary (Sinclair et al. 1987) based upon the Collins Birmingham University Language Database, there has been steadily mounting enthusiasm in tertiary education for the use of language databases or corpora for both research and teaching. A powerful means of exploring the language in a corpus (collection) of electronically stored texts is by using a concordancer. Concordancers are simple computer programs which can quickly analyse electronic texts to find any occurrence of a given word, part of a word, or phrase and display it within its immediate context. This paper will:

- suggest the potential of concordancers by providing a brief overview of how they have been used in tertiary education
- discuss the criteria necessary for developing corpora suitable for secondary school
- assess the feasibility of introducing concordancing into schools.

Concordancers in tertiary education

Valerie Pickard

So far, concordancers have been used in tertiary institutions for syllabus design and evaluation, materials preparation and interactive concordancing with students using the computers in the classroom. This section of the paper illustrates some of the benefits of using concordancers within these relatively privileged teaching and learning environments.

1. Syllabus Design and Evaluation

In the field of syllabus design concordancers have been used in two distinctly different ways. First, it is possible, using a corpus of learner English, to analyse the problems students have with the language. Various institutions are building large databanks or corpora of learner writing. In Europe, the Catholic University of Louvain is directing an international project to assemble and analyse one million words of writing by advanced learners of English from various language backgrounds (Granger, forthcoming). At the Hong Kong University of Science and Technology (HKUST), researchers are currently compiling and tagging a one-million-word corpus of the interlanguage of Chinese learners of English from the

Hong Kong Use of English examination and first year university assignments (Milton & Chowdhury, forthcoming). However such huge corpora, though potentially extremely valuable, are not essential to analyse the needs of a specific group of students. At The University of Hong Kong (HKU), Bruce (1991) analysed a 25,000-word corpus of student reports on the subject of Hong Kong's identity in order to investigate first year social science students' problems with logical connectors, reference and advanced level sentence structure. The results of his analysis were included in course material advising students on how to improve their English.

Instead of assessing students' problems with writing, the course designer may choose to use a concordancer to analyse the authentic target language in order to identify frequent words and phrases in addition to common syntactic patterns. For example, Ma (1993), teaching business writing at City Polytechnic of Hong Kong (CPHK), analyzed a corpus of 50 direct mail sales letters and observed the use of the imperative and modals in the various parts of the letters. His results will be of use to those developing courses with a business writing component.

Concordancers have also been used to evaluate existing course materials. Fang (1991) using a concordancer and word frequency analyzer he had designed himself was able to compare the vocabulary load in the five English Language textbooks commonly used in China. In three cases, the books compiled in China, he found that the vocabulary selection was generally unsystematic, whereas the other two textbooks, published abroad, adhered to certain principles of vocabulary selection and control.

Flowerdew (1991) found that ESP textbooks present a "distorted picture of syntax" when he used a concordancer to compare the language which first year biology students were actually exposed to with the language presented in ESP textbooks purporting to teach the language of the subject.

Pickard (1993a) used a concordancer to investigate the instances of refutation in academic writing and on finding very few examples, questioned the specific teaching of the language of refutation in EAP courses.

2. Materials Preparation (paper-based concordances)

Concordanced teaching materials can easily be prepared by the teacher and brought into the classroom. Students quickly get used to the appearance of concordanced exercises with their truncated sentences (see appendix 1 for an extract from a sample exercise). These materials are ideal for practising grammar and vocabulary items which students have difficulties with. At the same time they encourage an analytical and cooperative approach to learning. The teacher's role changes as students look for their own patterns and interesting points in authentic language: Johns (1991) calls this Data Driven Learning.

Concordanced exercises may take various forms, for example:

a. concordanced lines with the key word in context (KWIC) with guiding questions for discussion

b. gapfill (there may be one or more gaps for each word)

c. nonsense words instead of a gap

d. matching two halves of a sentence - before and after the key word

The above exercises may be used in a fairly structured, teacher-led discussion situation, or freer group discussion situation. Wu (1992) compares these two strategies which she used with two different groups of engineering students - those with 'A' level English were given more freedom than those with only 'O'level English. In addition, such exercises have been successfully exploited in a self-access system (Pickard, forthcoming).

3. Interactive Concordancing (students using a concordancer in the classroom)

When students are given the opportunity to use the actual concordancer program there is undoubtedly increased motivation as students are able to choose their own words or phrases to investigate. These choices may come as a surprise to teachers and will perhaps not have been anticipated by course book writers. For example, one pair of students using a concordancer in a class at HKU made their own gapfill exercise of words beginning with 'celebr' and found they knew 'celebration' but not 'celebrity'. In the same lesson a search for ***prod**** lead to a discussion about the difference between a 'productive manager' and a 'production manager.' In this way control of the learning is placed firmly in the students' hands. This by no means suggests that the teachers become redundant but that they are there to advise and facilitate; for instance, they might encourage students to sort the data in different ways, or suggest which dictionary or grammar books could be useful in order to help in the task of analysing the data.

Interactive concordancing has already been tried in some very different tertiary settings and for a variety of reasons. In England, Tribble (1991) has used a concordancer on EAP courses to analyze English writing across a variety of disciplines. In order to open students' eyes to the differences in style, he had them compare features such as noun phrases and post modification, verb phrases and grammatical structures of an economics textbook and an informal autobiography. After working on this, students were able to describe and identify various features of academic writing.

In the Middle East, working with students in Sultan Qaboos University, Oman, Stevens (1991) describes the steps he takes to introduce concordancing to first year university students with little or no computer experience or previous experience at discerning patterns in raw data.

At the University of Zimbabwe, teachers from the Communications Skills Centre have used interactive concordancing with students from a typically "teacher-dominated, rule-based learning system" to guide students to a better understanding of their new discourse communities by working on economics, geology and philosophy corpora (Mparutsa et al. 1991).

Closer to home for this Conference, Ma (1993) included an interactive concordancing component in a CPHK course teaching Computer Studies students how to write computer manuals. He observed the variety of learning strategies and techniques used by the students when using the computer program. A major recommendation he makes is that students should be made aware of the potential and limitations of the corpus.

Two short voluntary courses held at HKU trained students to use the Longman Mini Concordancer (Chandler 1989), provided them with teacher-developed concordanced exercises and concluded with students running their own mini-lessons or seminars using exercises they had developed themselves (Pickard 1993b).

Though mostly used so far for ESP or EAP at tertiary level, it is possible with the right corpus to have students in school English classes using a concordancer and concordanced-developed materials. The next section of the paper will consider what "the right corpus" for Hong Kong schools should comprise.

A Corpus For Hong Kong Schools

Kenneth Chan

If we assume that a specific rather than general corpus is most beneficial for concordancing in Hong Kong schools, one of the first questions to ask is "What sort of corpus is needed?"

In corpus design, 'representativeness' is important as it is in this (difficult to measure) quality that the usefulness of the corpus will reside. For the Hong Kong school situation, in order to tackle the above question we need to ask ourselves what kinds of writing our students are exposed to during their school lives. This includes exposure to writing outside school times, which could encompass, for example, any leisure reading that is not part of normal school work.

At school, a student's exposure to the written language is largely bound to the subject texts that are in use at different times; the texts may change at different stages of school life and so will the level of the language to some extent. Outside school, the non-subject related texts read by the student would include newspapers, magazines, story books, informal/formal correspondence, and instruction manuals of various descriptions.

1. Corpus Compilation

As indicated above, there are two main areas from which material for the corpus can be extracted, namely, (curriculum) subject-related texts, and non-subject related texts.

For subject-related texts, one way of building up this part of the corpus is to divide it into subject areas, i.e. English, Geography, History, etc., and extract portions from standard text books in use. These portions can be combined to form a larger archive of say, 50,000 words. Hence a small corpus on each subject can be built up and these corpora can, according to particular needs, be further and flexibly combined to form larger corpora. The advantage of this modular construction approach is that the subject corpora can be easily expanded, and integration of these into a larger, general or specific 'customized' corpus can be achieved simply by 'pick and mix.' For example a larger corpus of science can be made by merging smaller subject corpora of chemistry, biology, physics etc. Another advantage of this approach is that the substantial work involved in compiling can also be easily divided, as each subject corpus can be compiled independently from the others by teachers in a number of schools, provided that guidelines for selecting texts are well laid out.

For non-subject related texts the criteria for selection seem to be less clear-cut, and more research is still needed for a more accurate picture of this part of the corpus. Experience in working on a prototype corpus for HKUST students has shown that the following points are worth noting when compiling: the texts chosen should be:

- likely to be of general interest to students
- related to Hong Kong, or concerning affairs the students can identify with
- not highly analytical or over-technical
- not unduly restricted in their topicality (e.g. not about ephemeral news items)

2. Overcoming Difficulties

In the past, the process of extracting a source text, transcribing it using a wordprocessor and checking the accuracy of the copied version with the original took a great deal of time. Not only was this labour intensive but the technical difficulty was discouraging, especially to busy teachers in schools. However, now that computers are more technically advanced and affordable, as well as easier to use, and with more and more publications becoming available in electronic form, the work of compiling has become much more feasible.

Copyright concerns are never far away with this sort of work, and time and patience will be required in contacting and obtaining permission from publishers. Once complete, though, the product will be useful to many teachers and students in concordancing for a long time to come.

Concordancers in Secondary Schools

Janice Tibbetts

So far, this paper has outlined the teaching potential of concordancers and the criteria to follow in building up a corpus suited to work in secondary schools. All this effort would, however, be wasted where schools are concerned if teachers and students did not make use of the outcomes.

A concordancer is one of a number of CALL options which can be of use in secondary schools. Using computer concordancing shares many of the same advantages as other aspects of CALL. Learning with a computer has the advantage of being considered a "fun" way of learning, as opposed to what is rightly or wrongly regarded as the dreariness and boredom of traditional tools. Indeed, many an exercise that would be considered boring in a textbook becomes fascinating if done on a computer. But there are special advantages to concordancing that other CALL applications do not have.

1. Advantages

Perhaps the major advantage is that a concordancer can help in the empowerment of students in that it can be used as a tool for autonomous learning. It liberates students from teacher directed learning and brings them to the freedom of exploring language for themselves. For secondary school students, concordancers give opportunities to undertake serious research into areas of language they are having difficulties with or simply to play around and make discoveries for themselves. In Hong Kong this is especially useful at advanced level because of the fundamental change in the type of work required by advanced level examinations compared with the Hong Kong Certificate examination, which is taken at the end of fifth form.

As Mak (1990) has pointed out, this change in schooling marks a move from general and social English to a more academic type of language, and even an introduction to English for Occupational Purposes in that the examination has a section dealing with simple business English. Of course, it is not only in Hong Kong that this change occurs. There is traditionally a wide gap between the levels of examination at sixteen-plus and at eighteen plus. However, in Hong Kong, as in many parts of Asia and the developing world, this gap is extremely wide. The traditional language classroom has students working on the same materials at the same pace. There is little real communication, a lot of drilling and even the memorising of set vocabulary (often printed, helpfully, in a different colour in the reading passages of the textbook). Language is seen by most students and, sadly, by many teachers as only a **content** subject. The advanced level examination is, however, less concerned with content and more concerned with process. It is task-based, requiring candidates to handle large amounts of information at speed, to classify and make use of this information in order to identify and solve problems.

In this situation a concordancer used as a **resource** rather than a **course**, has real potential. It can lead to an increase in learner autonomy since students can look at what interests **them** or what **they** feel they need to find out. In doing this, they are receiving an early introduction to research skills. They have an opportunity to check their own use of language and to compare their personal language against authentic text. Wu (1992) found that students working on concordance exercises became actively involved in discussion to negotiate meaning and corrected each other whether the teacher was there or not. With the large classes of increasingly mixed abilities that exist in Hong Kong secondary schools this is clearly an advantage.

2. Constraints

In the long term the use of computers as cross-curricular learning tools is not merely feasible but inevitable. In the short term, however, the practical problems loom extremely large. These problems fall into four categories:- money, space, time and attitude. In Hong Kong the constraints are stronger than elsewhere, even than some places in the third world. (e.g. I found schools in Papua New Guinea were better designed and equipped.)

2.1 Money

One example will show the problem here. My own school's budget of $3,500 for forms six and seven covering about two hundred and thirty students is clearly inadequate, so the money will be spent on more books for teachers to use when setting examinations. If computers were available for language teaching, buying a concordancing program and license for multiple use would be feasible. But computers are themselves not available for reasons of space and time.

2.2 Space

Schools are overcrowded with too few, too small classrooms, leaving no space for a computer. Computer rooms may be located in rooms formerly judged unsuitable for teaching: these prove inadequate as both computer rooms and as language classrooms.

2.3 Time

Despite the disadvantages, the typical computer room is fully utilised every period of the school day for the teaching of computer studies, so the possibility of locating language classes there is, in any case, minimal.

2.4 Attitudes

Problems of attitude will be considered more fully. Experience and observation in schools can offer insights, with occasional surprises, into student and teacher attitudes in this area.

Student attitudes

Students' attitudes to concordancing or any use of computers in language are not a major problem given a good teacher and the right corpora. Most students seem to have a built-in aptitude for computers and find any work done on them highly motivating. The fact that they have more control over what they read and do with a concordancer is also a plus in their eyes.

When I took a party of sixth form students to Hong Kong University Practice Lab (a rather formidable term for a self access learning centre), I was able to observe student reactions to concordancing. They worked in pairs on the computers as this gave mutual support and the opportunity for discussion. It was eye-opening to see the different ways each student approached the situation and what each chose to do. All began by using the concordancer as a substitute dictionary, to find the meaning of words. They soon realised, however, that as a provider of meaning in the sense of definition the concordancer was inferior to a good dictionary. They then went on to investigate, or in some cases to discover, other aspects of language.

I would like to give just a few examples of the work done on this occasion. One student pair became interested in polysemy (the multiple related meanings of a word form) and in homonyms (identical word forms that carry unrelated meanings). They became very involved in this and found their explorations both useful and interesting. Another pair used the concordancer to study collocations and engaged other pairs in lively discussions on whether certain combinations of words were in fact collocations or not. They found much that interested and surprised them. Others concentrated on morphology, and enjoyed using the wild card marker to give a part of a word to see what forms the concordancer came up with. But possibly the most interesting point to come out of this very small scale trial was the discovery by students that the concordancer could show them different usages for words that they thought they knew. One pair, both girls, were initially somewhat dispirited by the exercise. They did not have the other students' enthusiasm for technology and were less adventurous in their use of the concordancer. Because there was no large corpus suitable for sixth form students, they found that the concordancer frequently reported that it did not have an example of the lexical item they requested. It was in a rather petulant spirit that they requested information on the word 'go'. Yet this experience convinced them of the value of the concordancer when it gave an example of the word used as a noun, as in 'have a go'. This was a totally new concept to them and generated much interest and excitement. The experiment, although on such a small scale, served to confirm my impression that students could benefit from using a concordancer.

Students' attitudes, then, pose no major obstacle to the introduction of concordancing in schools. The major problem lies elsewhere.

Teacher attitudes

In Hong Kong there is a surprisingly high degree of computer phobia among teachers. Computers are seen in terms of a subject to be taught rather than tools to be used. They are the sole responsibility of the computer studies teachers, who naturally use them to teach the computer studies syllabus. This requires students to learn basic programming skills. Once they have taken the examination most students do not use a computer again. Unless there is a computer at home, the majority of students do not acquire word processing skills or become familiar with database or spreadsheet programs. Teachers of other subjects have not yet awoken to the possibility of using computers as resources for teaching. Although many teachers are aware of the existence of learning packages for, say, geography and history, these are seen as having no relevance to teaching.

Teachers appear to fall into three main categories. The vast majority of secondary school teachers appear to belong to the first of these categories, that of computer illiterates. They do not use computers at all except to enter student test marks onto a computer disc. Even so, many of these teachers usually work on this task with a partner who enters the marks which they read aloud from a hand-written list. Such teachers express surprise and even admiration when they see others using a computer for other purposes. There is also, unfortunately, a tendency to look on such an occupation as 'playing' or 'showing off', instead of doing serious work such as marking books.

The second category consists of those who can use the computer as a kind of substitute typewriter, although most prefer the real thing. These teachers are interested in the potential for using computers to ease the onerous tasks of completing student reports and record cards. In some schools this is done already, but other schools are still investigating the possibility of introducing such a system. Some teachers in this category do use computers to produce examination papers or tests, and a very few produce teaching materials using word processing packages.

The third category consists of those who are completely at ease with computers and are familiar with word processing, database and spreadsheet packages. These programs are not seen as useful for teaching purposes, however, but for personal use or possibly for school administrative matters. Almost all the teachers in this category are teachers of computer studies.

Apart from their fear of computers themselves, many teachers also fear the loss of control that could result from the use of computers in the classroom. In Hong Kong, a very teacher-directed style of classroom management prevails. The teacher is the knower who imparts information, the expert who decides what is correct or acceptable. In order to maintain this role, language teachers here rely heavily on grammar rules and strict dictionary definitions to set the tone for 'correctness'. Lessons follow the text book closely and the 'teachers book' - which is usually nothing more than a set of answers to exercises - sets the standard for what is acceptable. Higgins (1988) pointed out that in many countries it is the students who fail to reach higher levels of tertiary education that become teachers and that language teachers in particular seem to have an inadequate grasp of their subject. This has led to the creation of the 'teacher proof' text book which actually discourages teachers from being innovative, and reinforces the concept of the one right answer, which is the one the teacher knows. The use of a concordancer with a large corpus of authentic text which may, and probably will, conflict with 'rules' of grammar, syntax and lexical definition, poses a frightening threat to a teacher's authority.

The senior staff in secondary schools are generally those who received their training before the advent of the personal computer revolution. As a result there is a general lack of knowledge about the potential for teaching with computers and a high degree of computer illiteracy. One principal confessed to never using the computer in his office. But, as he cheerfully pointed out, it gave a very modern and high tech impression to his office, which pleased parents. Within the Education Department, too, computers appear to be seen only as resources for the teaching of computer studies. Not enough are provided (in my school we have eleven computers for student use, although classes can have up to forty four students). There is no provision for the networking of computers, which is useful for self-access provision and for communicative language work on computers, but which is not necessary for learning how to write a program. No department has a budget for software provision and no materials have so far been provided to schools by the curriculum development division.

3. Feasibility

All this gives the impression that any attempt to introduce concordancing into schools is doomed to certain failure. But the picture is not so gloomy. Certainly students would welcome such a move and there is increasing interest in the possibilities among some teachers. If concordancing, and, indeed, other CALL packages are to be introduced successfully, the process must be carefully thought out.

It is probably not possible or even desirable to try to wrest control of the computer room from the computer studies staff. Their needs are also great and as we have already seen, the demands of the computer studies syllabus mean that the computers are heavily used for that course. Nor are computer phobic language teachers going to feel particularly happy about being shut in a room with equipment they fear, software they are unsure of, the pressure of adopting a new management style and a large, excited class of students who may, and probably do, know more about the equipment than the teacher does. In the authoritarian and hierarchical education system of Hong Kong the last point is the most terrifying.

The first step must be to reduce computer phobia among teachers. This will require commitment and care. Computers are already in schools; what is needed is training in use, organisation in arranging access, and commitment from senior management. Teachers are afraid of the machinery, in the same way that teachers in Europe once feared the language laboratory. There is the fear of damaging expensive machinery, of looking foolish if a student asks for help with the computer, of being unable to devise an appropriate methodology, and of being caught out by the superior knowledge of the machine - for students using concordancers may discover patterns which conflict with the teacher's rules. Even if they do not, teachers may be afraid that students will find something they don't know. The authentic text of the concordancer could be regarded as a rival to the teacher-expert.

Consequently, providing hardware and software will not be enough and both pre-service and in-service training is necessary. But before beginning to train teachers to use concordancing with students the ideal would be to make them feel comfortable with the machinery itself. A start can be made by actively encouraging the use of computers for administrative matters. Short courses in basic word processing skills should be offered in the training colleges for pre-service student teachers and computer-based self access material needs to be provided in the colleges. For in-service training, such courses should preferably be provided in schools. There are extra-mural courses already on offer but these demand a prior commitment to computers from teachers. The teachers must pay their own fees and give up their own time to do the courses. If administrators and the Education Department are seriously committed to training teachers to handle the machinery, then a policy of active encouragement must include allowing teachers the time and saving them the cost of being trained. Only when computers are regarded in the same light as typewriters will the possibility of concordancing approach reality.

The next step will be for teachers to be shown how computers can be used to produce materials for teaching and testing. Producing teaching materials is unlikely to have immediate appeal for many teachers here since few do produce their own materials, relying heavily on text books and past examination papers instead. (Cf. Adamson and Lee; this volume: editor's note.) But as a first step, demonstrating the value of computers in producing tests will certainly excite interest and it is here that a concordancer can prove its usefulness. (The feasibility of test production using a concordancer has already been investigated by Butler, 1991.) The discovery of the usefulness of concordancing in relieving the teacher of the burden of setting tests will be a major selling point. The setting of examinations in Hong Kong schools is a great headache because few teachers have expertise in writing tests. Another problem is that many teachers lack confidence in their own ability to use English, especially those teachers of other subjects who are drafted in to take one or two English classes to meet the shortage of English specialists. The usual solution is to simply copy tests and examinations from published materials. This has the disadvantage that most students are wise to this trick and many of them buy and use the same published materials for private study. In using a concordancer to produce tests, teachers will acquire familiarity and expertise which can then hopefully be transferred to the classroom or self-access centre. From tests it is not such a long step to move on to the production of practice material. In so using a concordancer the teacher will be building up the confidence she needs before using it with students.

Given sufficient encouragement and systematic support, teachers can be convinced of the educational and practical benefits of computers in general, and concordancers in particular, and these poentially valuable learning tools can then come to make their contributions in the school situation.

References

Bird, N., Harris, J. and Ingham, M. (eds.) (1993). *Language and Content.* Hong Kong: Government Printer.

Bruce, N. (1991). Concordancing analysis of students' end-of-year 'Hong Kong Identity' report. Unpublished report, English Centre, University of Hong Kong.

Butler, J. (1991). Cloze procedures and concordances: the advantages of discourse level authenticity in testing expectancy grammar. *System* 19(1/2):29-38.

Chandler, B. (1989). *Longman mini concordancer: text study software for teachers and students.* Harlow: Longman.

Fang, C.Y. (1991). COMPAID: A general-purpose package for literary studies. In J.C. Milton & K.K. Tong (eds.), (pp. 20-30).

Flowerdew, J. (1991). Corpus-based course design. In J.C. Milton & K. Tong (eds.), (pp. 31-43).

Flowerdew, L. & Tong, K.K. (eds.), (forthcoming). *Entering Text.* Hong Kong: The Hong Kong University of Science and Technology.

Granger, S. (forthcoming). New insights into the learner lexicon: a preliminary report from the International Corpus of English. In L. Flowerdew & K.K. Tong (eds.), (forthcoming).

Higgins, J. (1988). *Language, learners and computers.* Harlow: Longman.

Johns, T. (1991). Should you be persuaded: two examples of data-driven learning. *ELR Journal (New Edition),* 4:1-16.

Johns, T. (1991). From printout to handout: grammar and vocabulary teaching in the context of data-driven learning. *ELR Journal (New Edition)*, 4:27-46.

Ma, B. (1993). Small-corpora concordancing in ESL teaching and learning. *Hong Kong papers in linguistics and language teaching*, 16:11-30. (English Centre, The University of Hong Kong).

Ma, B. (forthcoming). Learning strategies in ESP classroom concordancing: an initial investigation into data-driven learning. In L. Flowerdew & K.K. Tong (eds.), (forthcoming).

Mak, L.Y. (1990). Integrating CALL into the H.K.A.L. curriculum. In V. Bickley (ed.), *Language use, language teaching and the curriculum* (pp. 577-594). Hong Kong: Institute of Language in Education, Hong Kong Education Department.

Milton, J.C. & Chowdhury, N. (forthcoming). An analysis of some of the characteristics of the writing of Chinese learners of English (Some problems and procedures in tagging the interlanguage of Chinese learners of English). In L. Flowerdew & K.K. Tong (eds.), (forthcoming).

Milton, J.C. & Tong, K. (eds.), (1991). *Text analysis in computer assisted language learning*. Hong Kong: The Hong Kong University of Science & Technology and the City Polytechnic of Hong Kong.

Mparutsa, C., Love A., & Morrison, A. (1991). Bringing concord into the ESP classroom. *ELR Journal (New Edition)*, 4:115-134.

Pickard, V. (1993a). Should we be teaching refutation? Concordanced evidence from the field of applied linguistics. In N. Bird, J. Harris & M. Ingham (eds), (pp. 387-407).

Pickard, V. (1993b). Lessons for concordancing. In N. Bird, J. Harris & M. Ingham (eds), (pp. 429-432).

Pickard, V. (forthcoming). Producing a concordanced self-access vocabulary package: some problems and solutions. In L. Flowerdew & K.K. Tong (eds.), (forthcoming).

Scott, M., & Johns T. (1993). *MicroConcord.* Oxford: Oxford University Press. (Manual written by Murison-Bowie, S.).

Sinclair, J.M., et al. (eds.) (1987). *Collins COBUILD English language dictionary.* London and Glasgow: Collins.

Stevens, V. (1991). Classroom concordancing: vocabulary materials derived from relevant authentic text. *English for specific purposes*, 10:35-46.

Tribble, C. & Jones, G. (1990). *Concordances in the classroom.* Harlow: Longman.

Tribble, C. (1991). Applications of stylistics in English language teaching. In J.C. Milton & K.K. Tong (eds.), (pp. 158-166).

Wu, H.H.M. (1992). Classroom concordancing: where can we go from here? In N. Bird & J. Harris (eds.), *QUILT AND QUILL: Achieving and maintaining quality in language teaching and learning* (pp. 166-183). Hong Kong: Government Printer.

WHAT MAKES AUTHENTIC MATERIALS DIFFERENT? THE CASE OF ENGLISH LANGUAGE MATERIALS FOR EDUCATIONAL TELEVISION

Arthur McNeill

1. Introduction

This paper describes an experiment in the development of English language educational television (ETV) programmes in Hong Kong. The experiment involved using samples of spontaneous English spoken by local Hong Kong personalities, all of whom spoke English as a second language. Traditionally, ETV programmes are carefully scripted and the language used is often intended to provide illustrations of particular language structures or functions. The artists who feature in the programmes are almost invariably native speakers of English, but are usually not professionally trained actors. This combination of didactic script and amateur artists has often resulted in programmes in which the use of English appears unnatural and the thematic content lacks real-world relevance. In an attempt to provide the target audience, Hong Kong secondary school pupils, with materials which might stimulate their interest in English and provide examples of English being used naturally by second language users in Hong Kong, it was decided to design a series of programmes based on (a) spontaneous (unscripted) spoken English and, (b) non-native speakers who use English to communicate as part of their everyday jobs in Hong Kong. This paper describes the materials, reports on teachers' reactions to them and compares the textual qualities of the materials with synthetic texts created for the same audience.

2. Authenticity and Language Teaching

The advantages and disadvantages of using authentic materials to teach a second language have been widely discussed. A number of researchers argue strongly for the motivating power of authentic materials, in particular their contribution to overcoming some of the cultural barriers to language learning (Kienbaum, Russell and Welty 1986, Nostrand 1989, Westphal 1986). Others (e.g. Beeching 1982) acknowledge that many authentic materials are simply too difficult for many language learners to deal with and can have a discouraging rather than a motivating effect on students. It is often difficult for teachers to find an appropriate pedagogical function for authentic materials, with the result that students do not always see the value in using them.

At this stage, it is probably useful to try to categorise the materials referred to in the study in terms of their claim to being considered authentic. It has proved difficult for applied linguists to agree about what exactly constitutes authenticity. A widely held view is that authentic materials are texts produced by native speakers for a non-pedagogical purpose (e.g. Bacon and Finnemann 1990). According to such a definition of authenticity, the materials used in the present study are not authentic since they are produced by non-native speakers of English. To restrict the definition of authenticity to texts produced by native speakers is obviously not acceptable in a second language context such as Hong Kong, where English texts are produced daily by Chinese native speaker civil servants, educationalists and businessmen as part of their official duties. In this connection, the definition put forward by Little et al. (1988:21) is more appropriate. They describe authentic language as "created to fulfil some social purpose in the language community in which it was produced."

While this definition is acceptable inasmuch as it accounts for the language of the texts used in the study, it does not say anything about the purpose for which the text is used or the way in which the texts are perceived by the target audience of secondary school pupils. Lynch (1982:11) warns of the complex semantics of "authenticity" and enumerates the synonyms for "authentic" used by a number applied linguists, to illustrate the different perceptions of the term which exist in the field. Widdowson (1979), in particular, provides a contrasting view to those mentioned above and argues that authenticity is not a quality of a text, but relates to the way in which a reader/listener responds to and interacts with a text. So far, the materials used in the present study have only been considered in terms of their authenticity as texts. The response to the texts as texts, rather than to tasks related to the texts, has been investigated by gathering teachers' opinions of the materials. Ultimately, the response of pupils to text-related tasks will have to looked into before any final assessment of the materials can be made. As Bacon and Finnemann (1990) point out, very little empirical research has been carried out to establish learners' cognitive and affective responses to authentic input. Their own research based on first-year Spanish students at two mid-western universities suggests that exposure to authentic input has a positive perceived effect on comprehension and learner satisfaction. However, the researchers point out that learners first need to be convinced of the value of dealing with authentic materials and recommend that exposure to them should start in the early stages of language learning and should be an important part of language instruction.

3. The Hong Kong "Authentic" ETV Materials

The programmes are intended for pupils in Form 2 of secondary school, most of whom are in the thirteen to fourteen age group. They form part of the Hong Kong Education Department's Educational Television Service's Bauhinia Series. (Senior producer: Yvonne Cheung, Director: Tina Wong, Language adviser: Arthur McNeill). The format of the materials is based on a structured interview, in which an interviewer guides an interviewee through a sequence of pre-prepared questions. The questions are intended to elicit expansive responses from the interviewee about aspects of his or her work. As far as possible they are designed in such a way that particular rhetorical patterns are prompted, such as descriptions, sequences of events, causes and effects, etc. Since the participants are not native speakers of English, the accuracy of their output cannot be guaranteed. However, in order to ensure fluency in the interviews, the interviewees were given the questions in advance and had the opportunity to chat informally to the interviewer before the recordings were made. The approach follows that used by the producers of the long-running BBC radio programme "Desert Island Discs." The general sketching in advance of the topics to be covered helps the interviewees to gather their ideas in preparation for the actual interview.

Since the programmes are intended to stimulate pupils' interest in learning English for use in Hong Kong, the choice of language informants was an important one. It was hoped that the pupils would be genuinely interested in what the interviewees had to say, both in terms of their "story" as successful Hong Kongers and in terms of the activities and processes which they describe. In the samples referred to in this paper the two interviewees are William Tang (Tang Chi Tak), a fashion designer, and Mr Mahmoud (Lam Mo Tak), who runs a music recording business. Not only are the individuals well known in Hong Kong, but their fields of activity are visually appealing and provide in the video plenty of visual clues to what is being discussed. The interviewer, Helen Yung (Yung Han Lung), was the first runner-up in the Miss Hong Kong contest in 1991. She is well known to the audience, speaks excellent English and is able to interact easily with the interviewees. Transcripts of four interview samples are reproduced in Appendix A.

4. Research Question

The general question which guided this first part of the study is as follows:

> How can the attractive features of authentic video be captured within a usable pedagogical framework in Hong Kong?

5. Teachers' Reactions

The four video extracts (Appendix A) were shown to 50 secondary school teachers of English, who completed a questionnaire which grouped their responses into four broad categories:

1. Features which might make the materials appealing to Hong Kong secondary school pupils.
2. Reactions to the high concentration of content words
3. The acceptability of Hong Kong English in teaching materials
4. Awareness of instances of incorrect English

The results of the questionnaire responses are summarised in Tables 1-4.

6. The Appealing Features

The teachers were first asked to note down any features which they considered might make the materials appealing to Hong Kong pupils. The most frequently mentioned features are summarised in Table 1.

Table 1

Features which make the programme appealing to HK secondary school pupils

	Feature	No. (n=50)	%
1.	Well known, interesting local personalities (success stories which might inspire pupils)	41	82
2.	Natural, spontaneous use of language	39	78
3.	A well known personality as interviewer	37	74
4.	Choice of topics	35	70
5.	Interviewer's clarification of new language/concepts	26	52
6.	Format easy to follow, yet not obviously an ETV programme	24	48

It is clear from the responses that the choice of local personalities was considered to have a very strong impact on the pupils' interest in the programmes. The fact that the two personalities were local success stories, engaged in international industries which prosper in Hong Kong, appeared to exert a special appeal. The majority of the teachers also considered that the natural use of language would make the materials popular with pupils. About half of the respondents stated that they found the interviewer's clarifications helpful in making the material accessible and comprehensible to the pupils. It might be argued that "authenticity" for these respondents is concerned with local relevance and reality, as well as naturally occurring English.

7. New Vocabulary

The second issue on which teachers' opinions were sought was the relatively high incidence of content words in the text, in particular the occurrence of some new, mostly technical vocabulary items. These occurred when the interviewees were describing aspects of their jobs. During the editing of the interviews, it was decided to include a number of new technical terms, since these arose naturally out of the situations described and could be regarded as essential. Examples of such words are "pattern cutter", "marketing people", "tape hiss", "high end" (of sound) and "modem". The teachers were asked to focus on the unfamiliar vocabulary in the text and to say whether the new words made the programmes too difficult or whether the new words were necessary to make the materials interesting. The responses are shown in Table 2.

Table 2

Teachers' reactions to the high concentration of content words

	Feature	**Agree**		**Disagree**	
		No.	**%**	**No.**	**%**
1.	The additional vocabulary makes programme too difficult for pupils to follow.	13	26	37	74
2.	Without these words, programme would lose authenticity and would be less interesting.	43	86	7	14

The majority agreed that the new words were necessary and contributed to the authentic flavour of the texts. It was considered that there would be a loss of interest if the new words were omitted. It was also pointed out that the meanings of the new words were made clear either by the context or the visuals or both.

8. Acceptability of Hong Kong English in Teaching Materials

One of the most frequently cited reasons for using recorded materials in language teaching is to provide learners with models of correct language use. Obviously, spontaneous texts spoken by non-native speakers are less suited as models. The teachers were asked to give their reactions to the use of non-native speakers in the videos and to give their reasons for supporting the use of non-native speakers, as well as giving reasons for any reservations. The responses are shown in Table 3.

82% of the respondents considered that the language spoken on the videos was "natural and spontaneous." This implies that, as far as they were concerned, it was authentic. Many pointed out that the samples of Hong Kong English were actually easier for their pupils to understand than English spoken by native speakers. It was felt that pupils' confidence in their own English would be boosted by the examples shown. It was pointed out by over half the teachers that the materials were best suited to listening practice activities, in which listening fluency rather than accuracy is the main focus. It was acknowledged that there was the possible danger that some teachers might use the materials as models, since this practice is so which accompany ETV programmes should discourage teachers from using these materials as models. However, the overwhelming response by the teachers was that the samples of Hong Kong were totally acceptable for inclusion in teaching videos.

Table 3

Use of Hong Kong English in teaching materials

	Feature	No. (n=50)	%
A.	Reasons for support		
1.	Natural and spontaneous	41	82
	Easier for students to understand than native speakers		
2.	Pace is relatively slow	33	66
3.	Inspires learners with confidence in their	29	58
4.	own English	23	46
5.	Helps to make students aware of typical Hong Kong errors	11	22
	B: Reasons for reservations:		
1.	Should be restricted to fluency work (listening practice)	27	54
2.	Need to avoid exposing learners to too many instances of inaccurate English	21	42
	Need to discourage teachers from using the		
3.	texts as models	19	38
	Need to maintain a variety of programme		
4.	types (e.g., dramatised stories, songs, etc. which involve more action)	12	24

9. Awareness of Errors

In the editing of the materials, an effort was made to omit any glaring errors in English. However, in the final version there are a few instances of incorrect grammar and pronunciation which could not have been removed without serious loss of content. In order to establish whether teachers were aware of these errors, they were asked to watch one of the excerpts (Excerpt 2) once without being asked to focus on the accuracy of the language. At the end of the first viewing they were asked to note down any instances of incorrect English they had noticed. They then watched the excerpt a second time and were asked to listen for errors. The responses are given in Table 4.

Table 4

Teachers' awareness of instances of incorrect English

		Not aware		Noticed 1 error		Noticed over 1 error	
		No.	%	No.	%	No.	%
1.	First viewing	35	70	8	16	7	14
2.	Second viewing	0	0	0	0	50	100

On the first viewing only 14% of the teachers were aware of more than one error, whereas all of the teachers were able to identify more than one error on second viewing. This suggests that the presence of occasional minor errors in a spoken text does not really interfere with comprehension. It might be argued further that if few teachers are aware of formal errors when they listen to a text for the first time, the chances of their pupils being aware of the errors is far less.

10. Textual Features

In order to gain a more objective picture of the characteristics of the materials, three measures were applied to Excerpt 4: lexical density, lexical variation and new word density. For sake of comparison, the same measures were applied to a published listening text intended for the same target group (Form 2) and widely used in Hong Kong schools. The listening text is not "authentic" by any stretch of the definition and was written for a pedagogical purpose: to provide listening comprehension practice. The listening text is reproduced in Appendix B. The results are contained in Table 5.

The lexical density is calculated by dividing the number of lexical (content) words by the total number of words and multiplying by 100. This gives us an indication of the relative distribution of lexical and grammar words in the text. In written texts there are generally more lexical words relative to grammar words than is the case in spoken texts. In Ure's (1971) analysis of 34 spoken and 30 written texts, it was concluded that the lexical density of spoken texts is generally under 40%, while that of written texts tends to be over 40% (37% - 57%). In the case of our "authentic" video sample text, the lexical density is 28%, while the synthetic commercially produced listening text for the same target group has a lexical density of 45%, which suggests that, in terms of the distribution of lexical and grammar words, it has more in common with a written than a spoken text.

Table 5

Some lexical comparisons of authentic video sample with published listening comprehension text for the same audience

	Sample authentic video text	**Published listening text***
Lexical density	28%	45%
Lexical variation (types/tokens)	21:64 (1:3)	83:160 (1:2)

*Chamberlain, J. *Communicating in English 2*, Hong Kong: Witman, 1987

The second measure, lexical variation, gives an indication of the type-token ratio in the text. This is calculated by dividing the number of separate words (types) by the total number of words (tokens). Lexical variation counts give us a rough measure of how many new items are introduced into a text as it unfolds (i.e. new to the text, not necessarily new to the learners). In the case of the video extract, the type-token ratio is approximately one to three, compared to one to two for the synthetic listening text, which suggests that the video text will be easier for a viewer to follow than the listening text is for a listener, provided not too many of the vocabulary items are unknown.

This brings us to the third measure, new word density, which is the % of words in a text which are new to the reader/listener. The teachers who took part in the survey were asked to estimate how many of the words in Extract 4 and the published listening text were new to Form 2 pupils. It is hard to get an accurate measure of new word density, since we cannot be completely sure that a word is new to pupils. The calculation is made more difficult by the presence of polysemous words (e.g. "high end" (of sound) "to mix (tracks)", "strings" and "brass" (in an orchestra). While pupils may know the words in their more general meaning, they may or may not be able to work out the sense in which they are used in the text. The results produced by the teachers indicate that the new words in the video extract account for less than 2% of the whole text, which represents a rate of occurrence of roughly one new word in every fifty known words. By contrast, the listening text had a new word density of 0%.

11. Conclusion

The presence of unfamiliar vocabulary does not, by itself, render a text difficult for a learner. Other textual features, such as lexical density and variation, need to be taken into account as well. As the teachers' survey shows, new vocabulary in a text can be a symptom of interesting content, which is necessary to sustain learners' interest. By contrast, the synthetic text mentioned above had no new language, which may account for its content-free flavour. Yet, according to the lexical density and variation measures, it was more difficult for the target group to process.

According to the results of the present study, it is possible to produce and use as teaching materials video programmes featuring Hong Kong personalities who speak English as a second language, provided they meet some basic criteria: (a) They need to involve real people whose lives and/or jobs are of genuine interest to the learners, (b) The individuals selected should be able to speak fluent, though not necessarily error-free English, (c) The texts should include a small proportion of new vocabulary, central to the topic, and (d) The visual context should provide an explanation of new language. However, before a more substantial framework can be put forward for the development of authentic video materials in Hong Kong, some empirical work needs to be carried out involving the most important personality in the whole equation: the learner.

References

Bacon, S.M., & Finnemann, M.D. (1990). A study of the attitudes, motives, and strategies of university foreign language students and their disposition to authentic oral and written input. *The Modern Language Journal*, 74:459-473.

Beeching, K. (1982). Authentic material 1. *British Journal of Language Teaching,* 20:17-20.

Kienbaum, B.E., Russell, A.J., & Welty, S. (1986). Communicative competence in foreign language learning with authentic materials. Final project report. (ERIC 275200).

Little, D., Devitt, S., & Singleton, D. (1988). *Authentic texts in foreign language teaching: theory and practice*. Dublin: Authentic Language Learning Resources.

Lynch, A.J. (1982). "Authenticity" in language teaching: some implications for the design of listening materials. *British Journal of Language Teaching,* 20:9-16.

Nostrand, H.L. (1989). Authentic texts and cultural authenticity: an editorial. *Modern Language Journal,* 73:49-52.

Ure, J. (1971). Lexical density and lexical differentiation. In G. Perren & J. Trim (eds.), *Applications of linguistics: papers from the 2nd AILA congress* (pp. 443-52). Cambridge: Cambridge University Press.

Westphal, G.F. (1986). On the teaching of culture in the foreign language curriculum. *Canadian Modern Language Review,* 43:87-93.

Widdowson, H. (1979). *Explorations in applied linguistics*. Oxford: Oxford University Press.

Appendix 1. Transcript of four interview samples

Excerpt 1 (Interviewer: Helen Yung, Interviewee: William Tang)

HY: Now, William, I know that you did not study fashion designing and so how did you start your career as a fashion designer?
WT: Well, I'd gone through my secondary school studying mathematics and sciences like most boys, and then at the university I studied economics and business.
HY: Economics and business?
WT: Yea, so towards the end of my fourth year, I told myself, you know, what really I want to do in the future. Should I do what my parents told me or should I do something that I believe myself. So I believe I should change my future into something that I believe. And I love drawing and sketching since my childhood. So I decided to leave Canada after I got my degree and I went to England for fashion design. That's how I started.

Excerpt 2 (Interviewer: Helen Yung, Interviewee: William Tang)

HY: What are the different stages involved in making clothes? Could you tell us what happens from start to finish?
WT: Very difficult. Well first of all, as a designer, you have to get your ideas together and put them onto a piece of paper - the sketching. OK, this is the first part.
HY: Sketching?
WT: And most people mistaken the first part. That's it. That's fashion design. It's not. It's only the step one, OK? From step one we move on to choosing materials. It's very very important whether the design is right for this material. You have to judge that, OK? So choosing material. Then after choosing material, you have to go on to pattern cutting.
HY: Pattern cutting?
WT: Yea, pattern cutting. So, you cut the paper and then you cut the material and then you sew it. So, this is the next step, to sew it. OK, after you sew it, the garment is right. Then you have to make sure everything is right - no mistake whatsoever. Then you have to carry on to manufacturing. But before manufacturing, you have to consult your marketing people or the salesmen whether you should produce or not.
HY: Consult marketing people and salesmen whether you should produce?
WT: Right and afterwards, you have to work together with your manufacturer. They have to give you time and space to produce all the garments. And after you produce them, you put them into the market to sell them.

HY: Into the market to sell?
WT: Yea, it's about that. And besides that, you have to do promotions.
HY: Promotions?
WT: Yea, advertisement.
HY: Advertisement?
WT: And all the money arrangements as well.
HY: Money arrangements? OK.
WT: Of course. It's really costly actually. So it's a really long procedure.
HY: It's a very long procedure.

Excerpt 3 (Interviewer: Helen Yung, Interviewee: Mr Mahmoud)

HY: Have you heard these songs before? Do you like them? Well, the composition and production of songs is a career that Mr Mahmoud has chosen since his secondary school days. Obviously, it's a very fascinating job for him, but let's hear more about this job from Mr Mahmoud himself. Hello there, Mahmoud. How are you?
MM: Fine.
HY: What are you doing?
MM: Well, I'm mixing a baby diaper commercial.
HY: Very interesting. In fact, all this equipment here is very interesting. What are they?
MM: Well, to start off with, in front of me is a mixing console. And what it does is it controls all the tracks that I have recorded and I can mix it together on it. This is a DAT. And what it stands for is "Digital Audio Tape". And this is what the tape looks like. See how small it is?
HY: Yea, it is quite small. What are some of the advantages of using DAT?
MM: Well, for one, it doesn't record music. All it records is 1s, 0s, 0s and 1s. So it's very clear. And it has no tape hiss.
HY: What is "tape hiss"?
MM: It means there's no tape noise, like the "ss" you hear on cassettes. And these are synthesizers and they're known as samplers because they sample sound, which means they record sounds digitally and they play back like a piano.
HY: Oh! Will you show me how?
MM: OK. Well,the top one I've got a trumpet sound going and it goes something like this.
HY: I recognise that. What about these two other ones?
MM: OK, well, this one is, eh, well, you listen first, OK? You tell me what it is.
HY: A car accident in the middle of rush hour!
MM: This is our recording room. And see how many instruments in here?
HY: There are a lot.
MM: It's pretty messy. And notice there's a mirror. Do you know why there's a mirror?
HY: No, why is there a mirror?
MM: Because the mirror reflects a lot of the high end of the sound and makes everything brighter.

HY: Can you give me an example?
MM: Yea, like "ss-ss-ss-ss-ss", so you can hear the "ss" a lot better. And this is the computer. I do most of my arrangements now on the computer. And what it does is when I play a note, it records exactly what I play and it shows on the computer. I'll show you. Let me take a seat. OK, now, this is what I just played. See? There's a line that goes through here.
HY: Yea, I see that.
MM: And that's exactly what I just played, so I can actually add something on it. See?
HY: Mm, very interesting.
MM: Very interesting. Yea.

Excerpt 4 (Interviewer: Helen Yung, Interviewee: Mr Mahmoud)

HY: So what do you usually do after composing the melody?
MM: That's a good one. What I did was I arranged the music after that, would go into a studio and record the music, supervise the vocals, which is the singer, and mix the music and the vocals after that. But nowadays, I use a team to do it. What I do is I record the music or the melody, so to speak, on the computer and I use a modem to send the material to an arranger. He lives far away. After sending it to him, I would wait about a day before calling him and listen to what he has done. Then I would amend whatever is necessary, over the 'phone, and he will send the data from his house to the studio and, after he does that, we will put it on tape, on the 24 track, and I will ask him to come in to supervise some live recordings, like guitars, or strings - violins, that is - or brass and, usually I'm not there, but after that we will ask the singer to come in and that's when I have to be there because the singer's rendition is going to be the most important thing on the tape because we are selling the singer. So, after that we do the mixing and the mixing session is actually as important as when the singer is in because that time we will have to put the music and the singer's rendition, the vocals, together and that determines our quality on the sound, so that's our procedure.

Appendix 2. Transcript of a commercially produced listening text.

V2: The job interview

Last summer, John applied for a temporary holiday job and was asked to attend an interview. The day of the interview came and John began to get ready in plenty of time. He put on his best jacket and trousers and tied his tie carefully in front of the mirror. Finally, he combed his hair. An hour before the interview he left the house and walked to the bus stop. The morning had been very hot and sunny but it was now very cloudy. John didn't have an

umbrella. It began to rain heavily. He stood under a bridge for a while but the rain didn't stop. Then he saw the bus coming. He ran as quickly as he could to the bus stop. The bus stopped and he got on. He was completely soaked. The traffic was bad and the bus didn't move very quickly. He was half an hour late when he finally arrived at the office for his interview.

V1: Now listen again and check your answers.
V2: The job interview

Source: Chamberlain, J. (1987) *Communicating in English 2*, Hong Kong: Witman.

RESOURCES FOR THE INDEPENDENT EFL LEARNER IN JAPAN

Stephen M. Ryan

1. Introduction

The concept of an autonomous learner studying without a text-book is a relatively new one in Japan. A heritage of Confucian ideas has been combined with strong centralised planning (Dore 1984) to make a transmission-model of education, involving teacher and text-book as the means of transmission, not only the traditional model but for many people the only model. From the traditional image of the priest in a temple-school surrounded by children eager for revealed truths from the sutras (Dore 1984) to the modern, surprisingly low-tech. classroom where the teacher expounds on the mysteries of the text-book while students listen and take notes (Rohlen 1983), the teacher and text-book are seen as the source of learning. Indeed, Miller (1982:45-48) has pointed out that this pattern is reproduced even on television, where educational programmes more often than not consist of a straight lecture to the camera in which viewers are referred to the text-book, "available from local book stores."

This cultural pattern is no less true for language education than it is for any other subject. However, in recent years there has been an explosion of resources in the country which could be used by language learners to improve their knowledge of English without the intermediaries of teacher or text-book. Indeed, although Japan is considered an EFL environment, in terms of the availability of resources for self-instruction it is coming more and more to resemble an ESL one.

This paper describes attempts to awaken EFL students to the possibilities of these resources and to encourage them to make use of them once they graduate from college and leave teachers and text-books behind. Although it gives an account of one particular tertiary situation, the paper also raises issues and implications of wider generalisability, at least within the educational and cultural contexts of Japanese society.

2. The Students

The students are all enroled in a 4-year engineering university in Osaka, Japan. They are on their way to careers as engineers and architects and their attention is focused on acquiring the technical and professional skills essential to their working lives.

As part of a programme of general education, the university provides them with 2 lessons of English language instruction a week. This comes after 6 years of English study in junior high school and high school, largely focused on passing the English test required for admission to the university. With the pressure of entrance exam. preparation removed and their main attention focused elsewhere, most students do not see English lessons as important or even relevant to their future plans, except for the thought that they may, at some point, take a trip abroad for pleasure.

In this judgement the students are sorely deceived. Graduates of the university have a high probability of working for a company (OIT Campus Guide 1993, p.46) which will send them to work abroad for an extended stay (a year or more). Whichever country they are sent to, it is likely that English will be the *lingua franca* between Japanese staff and their non-Japanese associates. Although such companies usually provide English lessons to those about to be sent abroad, they come amidst a host of other pressures and practical problems which severely limit the time and attention that can be devoted to mastering the language at that point.

Here, it was felt, was a strong case for encouraging learner-autonomy, so that, even in the first hectic years of a career, when it would be extremely difficult to make a commitment to attend English lessons regularly, graduates could improve (or at least maintain) their English by themselves.

3. The Course

It was decided to include, as part of the regular English course, an element that would alert the students to their probable future need of the language and encourage them to make use of the rich language-learning resources around them. The goal was not to transform them into autonomous learners during the period of the course but to show them what possibilities were available to them should they feel the need or desire to pursue their English studies at a later date.

The course consisted of modules, each focusing on a particular resource and containing information on obtaining and techniques for using the resource as well as an introduction to the principles that lie behind the techniques. A previous paper (Ryan, in press) outlines the methodology and procedures of the course. This paper will focus on the resources involved.

4. The Resources

After partially recovering from the initial shock of exoticism, first-time visitors to Japan are often surprised at the amount of English encountered in daily life (Barry 1992:121). Although this may suggest an abundance of English for our purposes, not all of this English is useful for self-instructed language-learning. It has been estimated that up to 10% of daily Japanese consists of recent loan-words from foreign languages, overwhelmingly English (Neustupny 1987:84-5). Words like bideo (video), kopiraitaa (copywriter) and shanpuu (shampoo) abound but the pronunciation of these words is often altered so radically in stretching them to fit the Japanese syllabic pattern (CVCV) that the Japanese forms are unrecognisable to native speakers of English. This is compounded by the fact that the original words are often abbreviated (hansuto [hun sto] is hunger strike and pasokon [perso com] is personal computer) and the user of the word usually does not know its origin. Furthermore, the meaning of the words is often altered too, so that manshon (mansion) is used for an apartment and dokutaa sutoppu (doctor stop) means "my doctor has told me to stop drinking."[1] Research into the origin of such words can be interesting and informative but is probably best not left to the unsupervised learner.

A further source of unhelpful English is the advertising and fashion industries which have taken to adorning their products with snippets of foreign languages, again with English predominating. T-shirts, bags, posters, pencil-cases, exercise books, even TV ads. bear strange, meaningless and often ungrammatical messages: "I love my puffing boy" "Lets enjoying piece in natural" "Live beer for live people." Fortunately most of this is regarded as decoration rather than a source of living English by most Japanese people so the amusement it provides for visiting foreigners is mainly harmless. However, when encouraging students to make use of the English in their environment, it is probably best to warn them off anything produced by the fashion or advertising industries.

Of the useful English resources that also abound in Japan, some are designed for instructional purposes, i.e. intentionally designed to teach English, and others are designed primarily to inform or entertain the resident foreign population. In the catalogue which follows, the resources are organised by the language skill they can best be used to develop. Multiple-use resources appear several times under different headings:

5. Catalogue of Resources Available to Independent EFL Learners in Japan

5.1 Reading

- Graded readers, such as *Oxford Bookworms* (OUP), *Longman Easystarts* or *Longman Originals* (Longman), *Variety Readers* (Cambridge) or the domestically produced *Ladder Series* (Yohan)
- Magazines and newspapers produced for learners of English (e.g. *Mini-world*[2])
- Novels (especially, books that have been turned into movies or are based on a movie)
- Movie scripts, like those in the Screenplay series (Fooin Creative Products)
- Technical manuals for particular hobbies (eg. computers, ham-radio)
- Foreign news-magazines (*Time, Newsweek*)
- Foreign newspapers (*USA Today, NY Times* - often available in college libraries)
- Japan's English language daily newspapers (available at station book-stores throughout the country, can also be delivered to homes)
- Foreign magazines - general interest
- Foreign magazines - particular hobbies or sports

- Locally-produced magazines for foreign residents (*Kansai Time Out* is a regional, monthly publication giving entertainment listings in English)
- Language-based computer games, designed for native speaker users (e.g. *The Carmen Sandiego* [Broderbund] series)
- Computer games designed to teach English (e.g. *Uncle Waldo's Will* [Intersoft])
- Letters from pen-friends

5.2 Writing

- Letters to pen-friends
- Letters to an English-language newspaper
- Articles to submit to magazine for English learners
- Poems, short-stories or video-scripts for fiction competitions (organised by local English-language newspapers and magazines)
- Guide book to the local area for foreign visitors/ residents
- Computer games designed to teach English
- Keep a diary/journal in English

5.3 Speaking

- Speech contests
- Recitation contests
- Exchange tapes with a pen-friend
- Button-hole native speakers, on trains, in English-language bookshops
- Make foreign friends, in bars or book-shops foreigners frequent
- Offer free Japanese lessons in exchange for free English lessons (for a text-book designed for this purpose, see Swanson & Swanson (1990))
- Join one of the many English conversation schools
- Travel abroad (especially alone, to increase chances of contact with locals)
- Home-stay abroad

5.4 Listening

- Tapes designed to develop listening skills in non-native speakers (readily available from all text-book publishers)
- TV and radio shows designed to teach English
- Radio stations intended for native-speakers (short wave, FEN)
- Radio stations that employ foreign DJs because English is trendy
- Satellite TV (one domestic channel broadcasts movies in English with subtitles in Japanese, another has regular news programmes in English from CNN, BBC, ABC with Japanese translation on another audio channel. Foreign signals [even Star] can be received but require expensive equipment)

- Sub-titled movies on TV with English sound-track (broadcast about once a month on terrestrial channels, several times a day on satellite)
- Movies on TV with multi-plex sound (multi-plex = original sound-track and Japanese translation broadcast on separate audio channels, receiver can select either channel) - broadcast several times a week
- Japanese TV news with English translation on a separate audio channel (broadcast almost daily)
- *Sesame Street* (broadcast in English with a Japanese explanation on a separate audio-channel)
- Videos of foreign movies (always with Japanese sub-titles)
- Closed-captioned videos of movies and situation comedies imported from the US (requires purchase of a decoder)
- Foreign movies in a movie-theatre (always with Japanese subtitles)
- Listen to foreign pop-songs (tapes/CDs often come with a paper giving the words in English and a Japanese translation)
- Exchange tapes with a pen-friend
- Button-hole native speakers
- Make foreign friends
- Offer free Japanese lessons in exchange for free English lessons
- Join one of the many English conversation schools
- Travel abroad
- Home-stay abroad

6. Initial Awareness of these Resources

A survey was conducted at the beginning of the course to find out to what extent students were aware of these resources. Several of the resources were listed and students were asked if they had ever made use of them for studying English and if they knew how to do so. Students were then asked to list other resources that they knew were available to them for studying English without a teacher.

Responses to the first part of the survey showed that less than 30% had ever made use of any of the resources listed, while the number who had made use of more than two of them was much smaller. Most students on most items responded that they did not know how to make use of this particular resource.

Students' suggestions for other resources they could make use of were somewhat fanciful. Some wrote of getting a foreign girlfriend or marrying a foreigner, others said they could speak English all day with their friends and one even suggested "be born as an English-speaking baby." Another group fell back on traditional study methods: copying words, studying a grammar book, text-book or dictionary, memorising place names (?) and talking with the teacher (a native speaker of English). It was clear from the responses that very few of the students had thought about the matter before and several of them had great difficulty with the concept of studying English without a teacher to tell them which materials to use.

Seven years' acquaintance with English, then, had done little to equip these students to make use of the resources around them to maintain their language ability once formal instruction ceased at the end of this course.

7. Consciousness Raising

The process of raising students' awareness of the resources had two goals: to make them aware that the resources were available and to convince them that they could make use of them without the supervision of a teacher.

The survey described above was the first step towards the first goal. It was intended not only as a means of finding out what the students knew about the resources but also as a way of having them begin to think about the possibilities for independent study of English that are to be found in their immediate environment.

After the survey, each module of the course was designed to present a different kind of material to the students. Information was provided about how to find the material, including such practicalities as how to arrange for newspapers to be delivered, exactly where in each bookshop English books are to be found and how to recognise that a native-speaker button-holed in a book-shop would like to end a conversation.

Care was taken to extract as much of this information as possible from the students themselves, in order to show them that they already knew or could easily find out most of the information. This sometimes consisted of simply eliciting the information and at other times it involved homework assignments like: "Find out when you can see TV news about Japan in English. What days? What times? What channels?" It was hoped that this would encourage students to reflect on other things they knew about which could be used as resources for language learning.

One of the requirements for the course was that students should keep a study-diary of occasions when they had studied English during the year. One section of the diary was set aside for further ideas on how they could study. By about the middle of the course most students were producing practical and realistic ideas.

However, without the courage to use them, producing information about resources for English study would be just an abstract intellectual exercise. For this reason, a lot of course-time was spent introducing students to exercises they could do by themselves, using the resources they were becoming aware of. After collecting information about multiplex news broadcasts on television, for example[3], students were guided through the following exercise:

1. Make a video of the day's news broadcast.
2. Identify the main news story.

3. Find an article in a Japanese language newspaper which explains the news story.
4. Predict English words that you will hear in the headline to the story.
5. Play the video of the first headline several times to see if your predictions were correct.
6. Find the same story in an English-language newspaper and predict further words you expect to hear in the headline.
7. Play the video again to check your new predictions.
8. Use a dictionary to check the meaning of any words you have successfully predicted that you do not understand.

It was emphasised to the students at each stage of this exercise that they were engaged in activities they could easily perform without the supervision of the teacher. At the end, they were pleasantly surprised to find they had understood the lead of a very complicated news story and learnt some new words. It was again emphasised to them that they were quite capable of achieving this feeling of satisfaction by themselves. By repeatedly performing exercises of this nature, accompanied by such overt propaganda for learner autonomy, students came to see that not having a teacher was no real hindrance to their exploitation of the language learning resources around them.

8. Selection Criteria

In the standard Japanese transmission-model of education, resources are selected by the teacher and imposed upon the students. The criteria on which selection decisions are based are assumed to be part of the teacher's professional competence, one of the things that distinguishes the teacher from the students. The result is that none of the students had any experience in selecting materials and it is likely that few of them had even thought about how such decisions are made.

In the course, students were introduced to three criteria for use in selecting texts: interest, intelligibility and feasibility.

8.1 Interest

Students were repeatedly reminded that one of the great joys of studying English without a teacher was that they could select resources which interested them personally: magazines connected with their own hobby, stories of a type they liked to read, news stories that caught their attention. From the reaction of the students, it seemed that few of them had considered the possibility that materials for study could be inherently interesting.

They were encouraged to apply techniques learnt in the course to resources which held some personal interest for them and to record their attempts in their study diaries. A marked reluctance to take on out-of-class assignments and pressure of work from their main subjects of study often prevented them from acting on this suggestion. However, the main reason for making the suggestion was to make them aware of the possibilities should time and motivation permit them to take their English studies further.

At another level, the idea of 'whatever interests you' was used to help them to make decisions about the specific language and language skills they wanted to focus on. The resources were presented as a smorgasbord from which they could choose according to taste ("I want to practise listening." "I need to improve my business English.") rather than feeling obliged to choose a balanced diet ("I haven't practised writing for a long time, so I suppose I'd better do some.")

8.2 Intelligibility

This criterion was presented mainly with reference to materials intended for learners of English. When resources designed for various levels of linguistic ability were available (as with graded readers or elementary, intermediate and advanced English courses on the television or radio) the different levels were explained and demonstrated.

Students were encouraged to estimate to which they could understand a text they had read or listened to. They were asked to give a round figure for the percentage of the text they felt they had understood (10%, 70%). This seemed strange to them at first but they gradually realised that they could improve their estimation by thinking comparatively: "I understood more of this text than the last one. Last time I said 40% so I suppose this must be more like 50%."

The purpose of encouraging such estimation was to overcome the students' feeling that they should not move on to the next level of difficulty before understanding completely texts at the current level. It was suggested to them that 80% comprehension[4] of a previously unseen/unheard text was an acceptable level of mastery before moving on to the next level of graded reader or TV English course. When initially selecting a level at which to begin, they were encouraged to find a level at which they could understand 40-60% of an unknown text.

8.3 Feasibility

It was not really possible to apply the criterion of intelligibility to materials produced for a native-speaker audience as such materials are not usually overtly graded by linguistic level. Instead, students were encouraged to select such materials based on the kind of tasks they felt could be achieved with the materials.

They were discouraged from seeking full mastery of resources designed for native speakers as this was thought to be a sure road to frustration. They were discouraged from this both overtly and by the nature of the exercises presented to them in the course. The exercises never focused on total mastery as a goal but approached each type of material with a view to learning something from it.

In one module, for example, ways of learning from local English language newspapers were presented and practised. These newspapers are written at a level of English far above that of the students. They can, however, be used to look for the spelling of proper nouns that feature in the news. A student seeing (from TV news or Japanese language newspaper) that a new Prime Minister had been elected in Greece could look in that day's Japan Times to find out how his or her name was spelt in English.

Thus, the third criterion for resource selection was the feasibility of using a particular resource to achieve a desired task.

9. Conclusion

It was never intended that the results of the course described in this paper should be measurable at the end of the year it took to teach. Evaluating the success of a course with goals that concern subsequent learner choices and behaviour is problematic. It is nonetheless already clear from students' comments that an awareness-raising course has a role to play in sensitising Japanese learners to the resources that are available for autonomous English study and to ways in which they can be exploited. What use the students make of this increased awareness in the future will be largely a product of circumstances beyond the control of the teacher. Be that as it may, this particular awareness-raising course has reportedly left students better equipped to prepare themselves to deal with future calls on their English ability than they were observed to be when they started it.

Notes

1 For a fuller listing of such words, see Webb (1988).
2 Produced by Mini-World Inc., Hirose Bldg. 1-13-12 Sekiguchi, Bunkyo-ku, Tokyo 112.
3 For a further, extended, example, see Ryan (in press).
4 These figures were based entirely on the author's intuition from his own experience as a language learner. They were certainly not presented as rules but as rough guidelines.

References

Barry, D. (1992). *Dave Barry does Japan.* New York: Random House.

Dore, R.P. (1984). *Education in Tokugawa Japan.* London: Athlone Press.

Miller, R.A. (1982). *Japan's modern myth: the language and beyond.* New York: Weatherhill.

Neustupny, J.V. (1987). *Communicating with the Japanese.* Tokyo: The Japan Times.

OIT campus guide. (1993). Osaka: Osaka Institute of Technology.

Rohlen, T.P. (1983). *Japan's high schools.* Berkeley: University of California Press.

Ryan, S.M. (In press). How to study English as content. *Memoirs of the Osaka Institute of Technology.* April 1994.

Swanson, P. & Swanson, Y. (1990). *If you teach me English, I'll teach you Japanese.* Tokyo: Japan Publications Trading Company.

Webb, J.H.M. (1988). *Katakana de no eigo no misu [Mistaken English in katakana].* Tokyo: The Japan Times.

ACTION RESEARCH CONTRIBUTES MORE TO TEACHING THAN JUST SOLVING DISCRETE PROBLEMS IN THE CLASSROOM

Annie Mueller

In this paper I would like to recount the experience of an action research (AR) group which worked together from autumn 1991 to spring 1993. The teachers involved were working at the English Centre of The University of Hong Kong (HKU), and came from a range of cultural and professional backgrounds. All taught a variety of English enhancement courses to first year students and had spent varying lengths of time at HKU.

The action research project led to significant concrete outcomes in individual classrooms and to changes in several of the courses taught at the English Centre - changes both in terms of underlying principles and in materials development. The project group shared a variety of professional concerns in addition to the research objective and by doing so explored numerous facets of teaching and its context. By providing such a forum, the AR project created a synergism in which all participants and the courses that were taught benefitted.

My interest in presenting this action research comes from the impression that while the AR project was under way, the participants consistently felt positive about the group, the meetings, the work it involved and the outcomes of the group's efforts. Sharing an analysis of why this was so will perhaps enable others to benefit from similar action research projects. Thus, the intention of this paper is to first give a short overview of the phenomenon of action research and then describe the action research project reported here. Finally a look at the *weak*, *strong* and *critical* versions of action research provides a framework on which to 'hang' the experience of this project in order to identify attributes which contributed to its success and which might be duplicated in general terms in other language teaching settings.

1. Overview of action research in education

Articles on or using statistical research or scientific research rarely begin by justifying their pedigree. However, the literature on action research often starts with a reference to the origins of this type of research which is perhaps an indication that AR is still not a universally known or applied practice. Since that may be the case, it is worthwhile to begin with some information about its development.

Modern educational research *per se* has been around for some time. Rusk (1932) cites a statement that the Teacher's Guild of Great Britain and Ireland made in 1888 regarding a study of **mental fatigue** (presumably among students). The study was thought to be "productive of much good and might help teachers to obtain ... influence to which they are entitled and which they do not at present possess" (Rusk 1932:16).

Rusk offers further indications of the value of educational research by and for teachers in giving the profession its due. He quotes a 1926 publication titled *Research for Teachers* - "Research will not only powerfully and rapidly develop the technique of teaching but will also react to vitalize and dignify the work of the individual teacher ... nothing would so effectively obtain for the teaching body the professional expertness and reputation for having it as the open-eyed, open minded scientific spirit of inquiry" (ibid: 69-70).

This indicates that already in the 1920s and 1930s there is an awareness of educational research and a concern that it should be undertaken for practical reasons, in order to understand, inform and improve classroom practices. In addition, research by teachers about teaching and learning is recognized as something that enhances professional self-esteem.

Rusk points out the distinction between **pure or scientific research** with its 2 steps:

1. a problem is selected
2. a careful and scholarly solution must be found

and **practical research** with 5 steps,

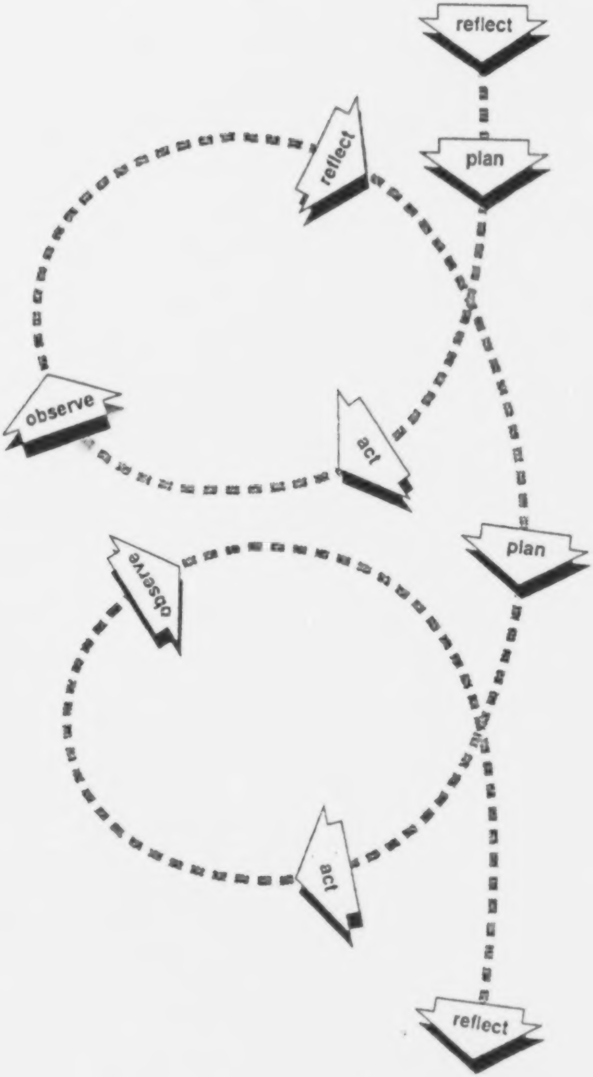
1. a going concern is studied
2. aspects are selected for investigation
3. a solution is generated in the *laboratory*
4. modifications are made so the solution can be put into *practice*
5. the solution is maintained by placing it into the organization to make it a permanent part of the system.

This is in some ways similar to a distinction being made today between traditional scientific research applied **to** education and action research **in** education (although action research does not generate its solutions in the laboratory). Traditional or scientific research is conducted to establish a truth which often remains as a written account - the privileged knowledge of a few. Practical research to which action research belongs, is intended to end in real life applications (Kember and Kelly, undated, p.1).

The term action research is generally credited to Kurt Lewin who in the 1940s and 1950s was one of several social scientists concerned with generating knowledge about a social system with a view to instigating real change (Elden and Chisholm 1993:121).

Figure 1

The Action Research Spiral



The standard representation of the action research cycle is a spiral. Figure 1 (reproduced with the authors' permission) is Kelly and Kember's (undated) diagram of Lewin's research cycle. The idea is not that some fact or truth is found, full stop. Rather, change is intended, a plan is made, action is taken and the results are observed. But that is not the end of it, for there is then reflection on the process undertaken so far, so that further or revised plans can be made, action taken and so on in the direction of the target of change. Succinctly put, "Knowledge without action is meaningless." (Elden and Chisholm 1993:122).

Action research was initially applied to social contexts such as minority groups within a dominant society. The idea was not to analyze situations objectively for the sake of increased knowledge, but to cause changes to occur which would improve social conditions. At about the same time there was similar concern for instigating action based on practical research in educational contexts. The dissatisfaction with traditional scientific and statistically-based research in education, already acknowledged by Rusk in 1932, was captured again by Stephen Corey in 1953 who wrote:

> We are convinced that the disposition to study as objectively as possible, the consequences of our own teaching is more likely to change and improve our practices than is reading about what someone else has discovered regarding the consequences of his teaching. The latter may be helpful. The former is almost certain to be. (cited by Oja and Smulyan 1989:4).

More recently such authors as Carr and Kemmis (1986), McNiff (1988), Nunan (1992) and others have extended the discussion and applications of AR in educational contexts. In Hong Kong, Kember and Kelly have produced the very useful booklet, *Action Research to Improve Teaching* which also has a worthwhile bibliography. The thrust of AR has been and is still primarily applied to social phenomena such as industrial and corporate settings, third world agricultural contexts, hospital environments and family counselling. Yet education is of course fundamentally a social activity and the language classroom in more recent decades has become an increasingly interactive and therefore social environment. Thus, AR for investigating the social phenomenon of the classroom and even whole curricula and institutions with a view to change has clear application for teaching and learning.

2. The action research project

The English enhancement teaching operation at HKU has been growing steadily since 1990. This has meant more students, more courses, more teachers, more committees and a change in management structure. This has further generated steadily increasing pressure on teachers' time - to teach, conduct tutorials, write and manage courses, carry out evaluations and engage in research. Some teachers thrive under these pressures. Most take a more restrained stance, particularly when faced with a choice of concentrating time on their students and classes, or, research. Students and classes regularly win out.

This was a situation in which action research was a perfect solution. As put by Allwright and Bailey (1991:199) " ... some of the energy teachers around the world currently put into collaborating on producing new syllabuses and new teaching materials in some institutions at least, could be rechannelled into longer-term but hopefully more ultimately collaborative work on classroom research." And this we were fortunate to be able to do.

The concept of action research was new to many of us at the English Centre in 1991 but its introduction was quite timely. The concept was appealing to teachers who wanted to do something with immediate and direct relevance to their teaching. Seven teachers became involved, about a quarter of the teaching staff. While this may have been serendipitous, Kowitz and Knudson (1980:45-46) maintain that the optimal size for problem solving groups is 5 to 7, as that number tends to insure a range of opinions while still allowing for each member to participate freely.

We spent several early meetings sharing our understanding of research in education and action research in general and in our local context. A list of relevant readings was compiled and put in a central place so all could access them when possible. We followed a process similar to that described in Allwright and Bailey *(loc. cit.)*

> ... groups of teachers ... share the initial task of identifying worthwhile issues to investigate by getting together to talk about the things that puzzle them about their classes ... parcel out the crucial but time-consuming task of reading through the literature to find out what has already been done and found out by other people ... [and then] come together again to design their investigation and to describe how it is to be implemented and monitored in their different classrooms. This team approach [is] also ... effective in analyzing and interpreting the resulting data.

Early on in the project, a single individual came forward as a leader who kept the group focused yet who was no more an expert than the rest. According to Kowitz and Knudson (1980:71), groups containing members who engage in high levels of orientation behaviour have a greater probability of achieving consensus. I think this is a significant factor in the success of all collaborative group work, and for us having such a group leader was a key factor.

We had fairly regular meetings for which a simple agenda, provided by the group leader, was generally linked to stages of the research cycle. Whereas one of the valuable attributes of our meetings was the latitude we enjoyed in our discussions, the fact of having an agenda also kept us sufficiently focused to go forward in our action research cycle. Taking notes and keeping minutes of meetings served the important function of formalizing the various group members' reports at different stages of the research cycle.

We followed the research cycle formulated by Kelly and Kember in their booklet. Allwright and Bailey (1991:44) add a gloss to the cycle:

1. identify an issue, interest or problem
2. seek knowledge
3. plan an action
4. implement the action
5. observe the action
6. reflect on your observation
7. revise the plan

From the group's experience, a preliminary step to the list above can be added and that is *the introduction to the idea of research by teachers* for their own purposes and the sharing of ideas and views on action research. This initial stage consolidated a group of like-minded people who were ready to cooperate in a shared endeavour.

The next stage was to *identify an issue, interest or problem*. The groups' members taught on different courses for different faculties but wanted to work as a group and all agreed to work on the same 'problem'. This was that although, over the years, more and more emphasis at our Centre was being placed on self-access work by students, the widespread perception was that the facilities were underutilized. Through discussion in the group, our attention was focused on learner training. The problem was eventually refined and stated as "To identify effective ways of motivating and training learners for self-access." This was something all group members felt could be investigated through an action plan leading to observable results.

At this point we gave ourselves a group identify and became the Self Access Action Research (SAAR) Group, not just an action research group. Fisher and Ellis (1990:46) list group identification as one of the factors in achieving cohesiveness and thus success in groups, and point out that a clear identity is valued by group members. In retrospect it was also important in giving us a recognition factor within our institution, and a certain status in relation to other established committees.

For the next step, *seeking knowledge*, we decided to validate our perception of the problem area by interviewing staff and distributing a questionnaire to get a more accurate understanding of the expectations teachers had themselves about self-access and what they felt about students' attitudes to SA. The survey results were compiled and from this information, reported in full in Martyn and Chan (1992), it was made clear that our perception of a need to impose learner training and motivation was accurate. What we needed to do next was formulate an action plan.

At this stage in the group's work we made several procedural decisions:

1. We would conduct our AR on a single course to maintain some degree of uniformity of participants

2. we would implement individual action plans within our overall objectives

3. We would record and monitor our action and observations using a common recording tool

4. We would meet regularly to share and discuss our progress or lack thereof.

(Martyn and Chan 1992:65)

The Action Plan form in Figure 2 proved a very useful tool in helping us to focus and consolidate our approaches, objectives and plans. It seems a very simple organizational device, but it was quite effective in empowering us to follow through on our action plans independently while at the same time allowing for comparability.

Details of the ***implementation of the action*** are fully presented in Martyn and Chan, (1992). In brief, seven different action plans were implemented individually. ***Observation*** was then undertaken by the individual teachers and experiences were shared in discussion by the whole group. The objectives and related action plans included among other plans ideas as diverse as self-access writing, student self-monitoring tools, and student perceptions and reactions to degrees of teacher control (for this last plan, see Nakhoul 1993).

The ***reflection*** stage of an AR project is a challenging part of the process. The difference between observing and reflecting in this context has to do with trying to determine what it all means - what the implications are and what the next steps should be. As Nixon writes, "Action Research prompts serious and often uncomfortable questions" (Nixon 1981:5). While each group member surely had individual and private reflections of value as well, the publicly shared reflections on the group's actions and observations contributed to concrete planning in terms of our objectives and in implementing new action plans.

Figure 2

SAAR Action Plan

Teacher:	EAS classes:
Semester 1 approach to self-access:	
SAAR objective: To identify effective ways of motivating and training learners for self-access. Specific objectives:	
Action: strategies, plans, etc.	
Observe: monitoring of students and self	
Comments: add on back of page if desired	

Devised by Elaine Martyn

Some features of the reflection process were:

1. *Planning for an in-house seminar on the group and its work.* This helped us to consolidate our thinking and to come to grips with organizing our observations in order to get them into a coherent and communicative format. It was also important to do this so that we could hear how other teachers in the same working situation responded to our observations and reflections. Obviously, their opinions and reactions would be valuable as an endorsement and would provide constructive criticism.

2. *Preparing for and participating in the SAAR group's meetings.* Knowing that we would be talking to the rest of the AR group about our actions and observations also had a crystallizing effect on our thinking. The meetings themselves gave opportunities to hear one's own thoughts aloud and to have ideas questioned or extended by others.

3. *The monitoring tool.* The action plan/record form (figure 2) was not only useful for observation but also for retrospection or reflection and planning forward. It provided us with topic areas about which we all could contribute based on our individual experiences.

4. *The consideration of other courses.* Because our topic - self-access - potentially involved other courses besides those on which we taught, we were compelled to look at these other courses to a certain extent. This widening of our frame of reference became a process of cross-fertilization - hearing about ideas in other courses, hearing what other teachers were doing and planning, what worked and what did not. We also became concerned throughout the project with *teacher training in self-access* as well, and this gave us an additional reason to look at what teachers were doing or expected to be doing in other courses.

5. *Liaison with course coordinators or managers.* We invited staff responsible for course design and coordination to our later meetings to share our reflections with them. This liaison effort enabled some of the group's work to become institutionalized in course materials, as well as giving the coordinators and course materials writers stronger justifications for pursuing self-access as a required component of a course.

The final stage of the cycle - *revising the plan* - included:

1. *A workshop for teachers* on self-access facilities so that they would be more confident in encouraging and guiding students toward independent learning.

2. *Raising teacher awareness* about the learning styles of our students and the wider context of the university. Both of these factors clearly have great effect on what students are willing and capable of achieving in our English enhancement classes. This has caused us to reflect on and sometimes adjust our expectations.

3. *Revising course materials* which further encourage both teachers and students to use self-access as a viable and productive feature of their coursework.

4. Far more concentration on *developing the self-access facilities* both conceptually and in terms of materials and consulting services for students.

5. *Exchanging and clarifying views on our teaching.* Hearing what 'old' teachers had to say about courses and what 'new' teachers had to say gave us a much fuller picture of teaching practices at our Centre. Through reflection we were able to establish elements of our teaching situation which were being unduly taken for granted as well as to clarify objectives and expectations which were far from transparent for new teachers. This had implications for teacher's notes in materials and in improving the orientation process for new teachers.

3. Weak, strong and critical versions of action research

Finally, the *weak*, *strong* and *critical* versions of action research provide a framework with which to analyze the process. (The terminology can be challenged but will serve the purpose.)

Peters and Robinson (1984) have discussed action research as a methodology. After having surveyed a number of action researchers, they determined that two versions exist.

The *weak* version of action research is a basic problem solving strategy or methodology and may not be based on a particular philosophical or social science approach. This approach typically involves an independent or outside expert and a client group who wish to improve an existing practice (ibid. 121). Candlin (1993), in the course of a presentation at Hong Kong Polytechnic, referred to this version of action research in education as teachers dealing with puzzles they face in the classroom, indicating that the action plans and observation and reflection are contained within that context.

Peters and Robinson's second version is called *strong* AR. In this version, researchers do not rely on the independent role of an outside expert and instead focus on the equal participation of group members in all aspects of the research process from initial problem formulation to the implementation of strategies (ibid. 121). Candlin's interpretation of the *strong* version of AR was that it is collaborative, multi-functional and deals with problem identification, teacher development and institutional change.

It seems that, in terms of this account, it is at the level of problem identification where the AR is determined to be either in the category of *weak* or *strong*. Teachers' initial attempts at AR may be limited to the *weak* version until they develop the confidence and methodology to pursue a stronger and more independent type of research. One must begin at the beginning, however. Implementing *strong*

AR also seems to be dependent upon the willingness of a group to pursue a problem which involves institutional change as well as classroom concerns.

The *critical* version of AR is described by Elliot (1980: 321-322) as a research theory that would explain the ways in which teaching is constrained by factors operating outside the classroom in its institutional, social and political context. Further, there is the implication of increasing the professional autonomy of teachers through a critical version of AR. (This idea of professional autonomy in education, we noted, was already a concern back in 1926.) Carr and Kemmis (1986:130) comment on the emancipation phenomenon associated with action research. Their discussion concerns the development of AR as a reaction to a positivist scientific approach to research that assumed an objective reality and a detached observer. A *critical* research theory then has the central task of emancipating people through their own understandings and actions.

By way of concluding, I would like to try and establish what our SAAR project accomplished in terms of the *weak*, *strong* and *critical* versions of AR. What did we do that made us feel so positive about our efforts, and what might other AR groups consider as successful procedures or strategies?

In terms of *weak* AR we agreed to develop individual action plans for our separate classes within the larger umbrella of learner training. I think that flexibility in our group dynamic was a necessary and sensitive component of our overall affective success. Our language classrooms are our domains, where we work in our own ways to build rapport and confidence with our students; devising action plans which reflected our personal expectations enabled us to work effectively according to our own individual styles.

"Weak" may not be the best term to convey a focus on the individual classroom. Our individual classroom action plans and observations formed the nucleus of all the subsequent discussions that went on as well as the substance of several publications resulting from our AR. It is in the classroom that the fundamental action occurs and the primary observation of change takes place. If a teacher or group of teachers are concerned with issues in their own classrooms only, they are nevertheless instigating real change based on the AR cycle. Monitoring the plan, action, observation and reflection stages is a process we found to be professionally satisfying.

In our SAAR meetings, sharing the confusion and concerns over our individual classroom problems as well as the pride and satisfaction of positive changes in a genuinely interested forum was a very satisfying experience both personally and professionally. In discussing the benefits of peer networks in teaching contexts which in fact our SAAR group was, Sithamparam and Dhamotharan (1992:12) state that the supportive atmosphere of such groups is necessary if teachers are to attempt innovation that will entail personal readjustments as well as gradual accommodation in thinking and practice.

It was the synergistic synthesis of the individuals in the AR group which made the project strong. One of the group's members wrote of the project in an in-house newsletter:

> Both the synthesis of our different insights and the supportive but analytic criticism offered have been important to the continuing 'reflection' which takes place within the group. Moreover, the act of sharing our perceptions has been inspiring and reenergising in itself. Group action research leads to a synergistic state which maximizes the benefits of professional development for the individual, group and institution.
>
> (Nakhoul, in-house communication).

In terms of *strong* AR, we were not only sharing experiences but also collaborating. For some researchers, participatory and/or collaborative research suggests that outside researchers with experience come into the teaching situation and guide teachers in their action research. Examples of this are in Oja and Smulyan's *Collaborative Action Research* and Whyte's *Participatory Action Research*. While this may be helpful in some contexts, we found that we were able to conduct our own research program with support from the relevant literature and from our own motivation.

Our group's interests drove us to consider numerous aspects of our teaching context as we realized we were opening a window on our work which had not been looked through before. Out of the group's observations and reflections we gained insights into our particular teaching and learning world which we were then able to share with our whole Centre. We developed our understanding of what the Centre was expecting of students and how we could or could not help students realize those expectations. Our AR had value not only for our classroom practices but for course development and for language learning resources outside the classroom. The group was able to offer advice with confidence and make suggestions to staff about ways they might take advantage of our experiences and observations and reflections.

Further strengths of collaboration in group work are mentioned by Barker et al. (1987:10-11) in a book entitled *Groups in Process*:

* groups are more efficient than individuals in the recall of information

* correct or accurate responses are made more often by groups than by individuals

* fewer errors in judgement will be made by groups than by individuals

To the extent that we may have engaged in a *critical* version of action research, I would like to speculate that it was the sense of professional autonomy we achieved which made us most satisfied. The teachers at the Centre had inherited the idea and movement toward increasing self-access in our teaching context. To say this had been imposed is too strong, but there was some uncertainty about what was meant by self-access and there had been little inquiry about it except superficially. Expectations on the part of students and teachers were not initially clear.

What the project group did was take control of the issue by following through in a complete action research cycle. The planning, actions, observations and reflection caused us to look in detail at assumptions and misconceptions about ourselves and our students. Since our overall topic was learner training we had to consider our training and our students' training prior to being in our Centre. We had to consider the expectations we had for students, what expectations they had for our teaching and how we compared to other teaching contexts at the university. We had to have a hard look at the realities of the total educational environment of the university and even at what students would expect to do upon graduation. The more we knew about the situation the more we were able to direct it in ways we had determined would be most effective, thus gaining a measure of freedom. We took control, which was satisfying personally and professionally.

4. Closing remarks

Schecter and Ramirez (1992:205) have recommended that teacher researchers share their experiences for two reasons:

1. Close observation of the workings of such groups can reveal the significance and value of teacher research as perceived by teachers and it can also help pivotal actors in teacher-research projects to take appropriate actions.

2. The study of a variety of teacher-research groups that differ in configuration and goals can provide a needed understanding of the spectra along which different groups may coexist and of the outcomes yielded by various combinations of points along these spectra.

I hope that this paper will have contributed somewhat to these two goals.

References

Allwright, D. & Bailey, K.M. (1991). *Focus on the language classroom.* Cambridge: Cambridge University Press.

Barker, L.L., Wahlens, K.J., Watson, K.W. & Kibler, R.J. (1987). *Groups in process.* New Jersey: Prentice Hall.

Candlin, C.N. (1993). Seminar on action research. Hong Kong: Hong Kong Polytechnic.

Carr, W. & Kemmis, S. (1986). *Becoming critical.* London: The Falcon Press.

Chisholm, R.F. & Elden, M. (1993). Features of emerging action research. *Human Relations,* 46(2):275-297.

Elden, M. & Chisholm, R.F. (1993). Emerging varieties of action research: introduction to the special issue. *Human Relations,* 46(2):121-141.

Elliott, J. (1980). Implications of classroom research for professional development. In E. Hoyle & J. Megarry (eds.), *World yearbook of education 1980: professional development of teachers.* London: Kogan Page.

Fisher, B.A. & Ellis, D.G. (1990). *Small group decision making.* New York: McGraw Hill.

Kember, D. & Kelly, M. (undated). *Using action research to improve teaching.* English Teaching Unit, Hong Kong Polytechnic and the English Teaching Centre, City Polytechnic of Hong Kong.

Kowitz, A.C. & Knutson, T.T. (1980). *Decision making in small groups.* Boston, MA.: Allyn & Bacon.

Martyn, E. & Chan, N.Y. (1992). Self-access action research: a progress report. *Hong Kong papers in linguistics and language teaching,* 15:59-68. Hong Kong: The English Centre, Hong Kong University.

McNiff, J. (1988). *Action research: principles and practice.* London: MacMillan Education.

Nakhoul, L. (1993). Letting go: preparing teachers and students for learner independence. In J. Edge & K. Richards (eds.), *Teachers develop teachers' research* (pp. 147-155). Oxford: Heinemann.

Nixon, J. (1981). *A teacher's guide to action research.* London: Grant McIntyre.

Nunan, D.C. (ed.), (1992). *Collaborative language learning and teaching.* Cambridge: Cambridge University Press.

Oja, S.N. & Smulyan, L. (1989). *Collaborative action research; a developmental approach.* London: The Falmer Press.

Peters, M. & Robinson, V. (1984) The origins and status of action research. *The Journal of Applied Behavioral Science,* 20(2):113-124.

Rusk, R.R. (1932). *Research in education.* London: University of London Press.

Schecter, S.R. & Ramirez, R. (1992). A Teacher-research group in action. In D. Nunan (ed.), (pp. 192-207).

Sithanparan, S. & Dhamtharan, M. (1992). Peer networking towards self-direction in teacher development. *English Teaching Forum,* 30(1):12-15.

Whyte, W.F. (ed.), (1991). *Participatory action research.* California: Sage Publications.

COMPUTER-MEDIATED COMMUNICATION AND TEACHER EDUCATION: THE CASE OF *TELENEX*

David Coniam, Sima Sengupta, Amy B.M. Tsui and Wu Kam-yin

1. The Conception of *TeleNex*

Amy B.M. Tsui

1.1 Introduction

In the last ten years, there has been an increasing exploitation of telecommunications for educational purposes. Many computer networks have been set up to enable students to communicate with each other inside and outside the country. Networks have also been set up to link schools with tertiary institutions so that students can access the university libraries and communicate with university staff (see for example Meadows 1992). The application of telecommunications in teacher education, however, is very recent (see also Davis 1991). This paper describes a computer network *TeleNex* (Teachers of English Language Education Nexus) set up by TELEC (Teachers of English Language Education Centre) at The University of Hong Kong to provide professional support to English language teachers in secondary schools in Hong Kong. This section outlines the conception of this network, and presents its database and communication components. Later parts of the paper offer more detailed discussion of the issues involved in the design and organization of the three databases that comprise the present database component.

1.2 Computer network and professional support for ESL teachers

In recent years, there has been a constant shortage of teachers in all subjects, both as a result of the introduction of compulsory education in Hong Kong in 1978 and as a result of an increasing number of people emigrating with the change of sovereignty in 1997 drawing near. The problem of teacher shortage in ESL is further aggravated by the fact that ESL takes up the highest number of teaching hours in the curriculum, hence requiring more teachers, and the fact that graduates who are proficient in English tend to be lured away by business with lucrative offers because they are very much in demand. This results in an ESL teaching force that is largely composed of teachers who were not subject trained. In a survey of teachers in Hong Kong conducted by the Education Department in 1991, it was found that among the 3700 graduate teachers of English, only 27% are subject trained, that is, they were English majors in their undergraduate studies, and only 21% are both subject and professionally trained (author's analysis). The picture is even more disconcerting when one looks at the whole secondary English language teaching population, which is 5240. The percentage of teachers who are subject trained drops from 27% to 18.9 %, and the percentage of teachers who are both subject and professionally trained drops from 21% to 14.2%. The percentage of English teachers who have had initial and refresher training is 7.4% and the percentage of graduate English teachers who have had initial and refresher training comes to as low as 2.3% (see Tsui 1993).

In view of the academic and professional background of English language teachers, it was felt that apart from the conventional way of providing in-service initial and refresher training courses, there should be a way in which teachers can obtain professional support wherever and whenever they need it. There should also be a way in which teachers can give each other professional (and emotional) support. Drawing on the experience of CATNET, which is an electronic bulletin board system set up for computer teachers in schools in Hong Kong where members join on a voluntary basis from their home computers, it was thought that a similar network could be set up for English language teachers. However, the experience of CATNET shows that discussion among teachers must be sustained by a discussion leader, otherwise it will die off very quickly (see Ip 1989). Therefore, rather than replicating what other networks have been doing, it was felt that the network to be set up should be monitored and supported by academic staff at the university. It was also thought that the network could be further exploited to provide resources for teachers to help them in their everyday teaching and to enhance teachers' knowledge base by providing them with some basic information about how the English language works. The idea was put forward to Hong Kong Telecom Foundation who agreed not only to provide funding for the project but also the hardware and the communication lines. TELEC was established in June 1993, with the full staff , consisting of the equivalent of three full-time academic staff, one computer and two clerical staff in place in Sept 1993. The academic staff are ESL teacher educators and experienced ESL teachers (hereafter referred to as the TELEC team). By the end of November 1993, 15 schools were linked up with the Centre with over 130 teachers registered as participating teachers.

1.3 Components of *TeleNex*

Most computer networks enable users to access information and/or to communicate with each other via electronic mail to share ideas and to collaborate on projects. *TeleNex* differs from these networks in that it is monitored and supported by a team of ESL teachers and ESL teacher educators, as mentioned above. Apart from enabling teachers to talk to each other via e-mail, to collaborate on projects and to seek advice from the Centre staff, it also provides several databases which are specially designed and written by the team to meet the needs of the teachers. The following diagram summarizes its components and their functions.

As we can see from Figure 1, there are two components to the network: database and communication. The database component consists of three databases. The test bank database provides a bank of validated items from which teachers can draw items and compile test papers. It also encourages teachers to send in test papers that they have produced themselves which, after validation and selection,

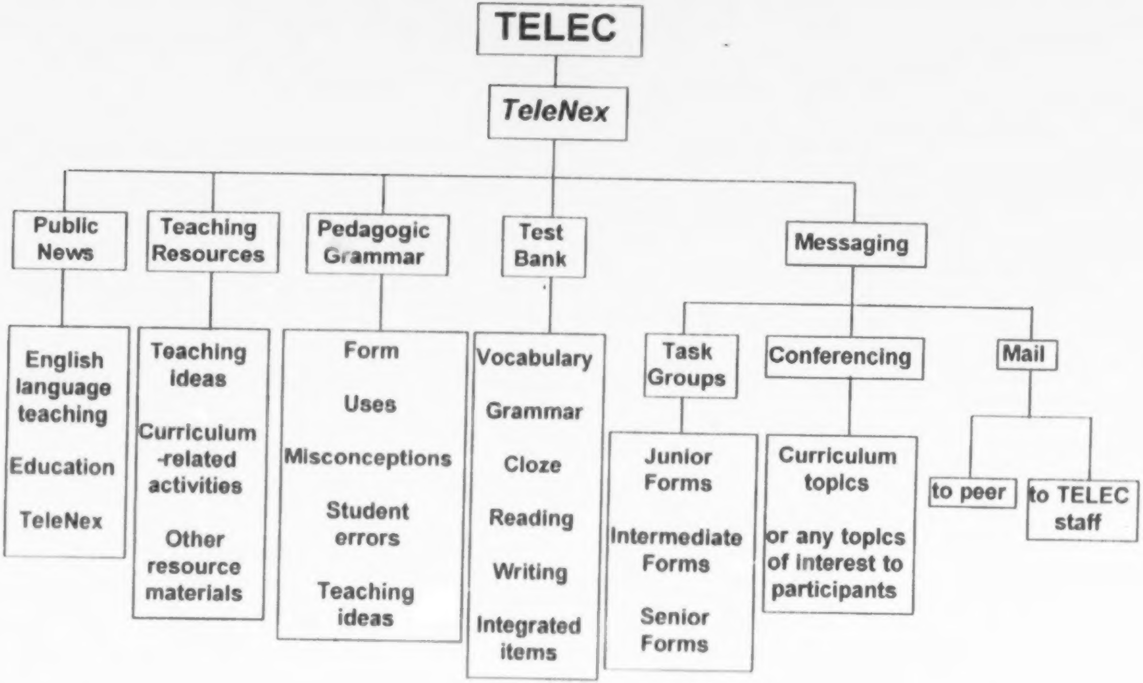

TELEC
TeleNex
Public News
Teaching Resources
Pedagogic Grammar
Test Bank
Messaging
English language teaching
Education
TeleNex
Teaching ideas
Curriculum -related activities
Other resource materials
Form
Uses
Misconceptions
Student errors
Teaching ideas
Vocabulary
Grammar
Cloze
Reading
Writing
Integrated items
Task Groups
Conferencing
Mail
Junior Forms
Intermediate Forms
Senior Forms
Curriculum topics
or any topics of interest to participants
to peer
to TELEC staff

Figure 1

will be stored in the bank for teachers to use. (See Section 2 for a detailed discussion of how this database works.) The teaching ideas database aims to help diagramteachers in their everyday teaching. Teaching ideas and the rationale behind the ideas are provided as well as some classroom materials and activities that teachers can download and use in the classroom. (See Section 3 for a detailed discussion of the design of this database.) The grammar database aims to enhance the subject knowledge of ESL teachers by providing some basic information about how the English language works, what the common misconceptions and how best to teach them. (See Section 4 for detailed discussion of the design and organization of this database.) Finally, the news database provides information about current developments in education and English language teaching, which is an important part of the professional development of teachers. It also keeps participating teachers informed of the recent developments in *TeleNex*.

1.4 Features of *TeleNex*

There are two outstanding features of *TeleNex*: it is both interactive and collaborative. Databases are usually for information retrieval only. The databases of *TeleNex*, however, are interactive. Teachers can send in their comments and questions about the database as they are going through it. Upon receiving comments and questions, the Centre staff will make the necessary revisions in the databases. This enables the databases to undergo constant revision so that they will better meet the needs of teachers. The databases are also collaborative. The resource materials and activities are initially written by the TELEC team. However, teachers are encouraged to send in materials that they have tried out in their classrooms and have worked well. Both the grammar and the teaching ideas databases now contain contributions from participating teachers, most of which have been tried out and have proved to work well.

The communication component consists of three parts: mail, conferences and task groups. The mail part allows teachers to send questions to the TELEC team for advice or for answers, or to send individual mail to each other. The conference part allows teachers to exchange views, ideas and even frustrations. This kind of sharing and mutual support is very important for the morale and the professional development of teachers. As the director of one of the largest collaborative networks in the US points out, "The network makes teachers feel empowered. When they've had a discouraging day, they have someone they can talk to." (Mangan 1992). The task group allows teachers to collaborate on curriculum projects and address problems in teaching. These task groups are organized by teachers themselves. Any teacher can start a task group if he/she wants to and invite other teachers to join. For example, a Form Five teacher may be worried about preparing students for the new public examination on speaking skills which will be introduced in April 1994. He/She can start a task group to address this problem. Teachers can share ideas, materials, or even work on some oral packages to prepare students for the examination. These task groups will be monitored by the TELEC team to ensure that the discussion is sustained and that an outcome is achieved. Task groups will be closed down when they have accomplished their tasks. Hence, instead of the teacher just looking to the TELEC team for support, the teachers giving each other support which is very important in empowering teachers and helping them to develop professionally. In other words, *TeleNex* is a network supported collaboratively by teachers and teacher educators.

In addition to providing service to teachers, there is a further research element in *TeleNex* which is built into the computer programme from the very beginning. The software allows us to keep track of the patterns of use. This includes aspects like which screens are often browsed by teachers and which ones are never browsed, what are the peak hours of use, what are the patterns of interaction, which task groups flourish and which die very quickly. All of this is valuable information which helps the TELEC team to understand the teachers better and to improve the network.

By inviting teachers to send in ideas, materials and test papers that have been tried out in classrooms, and by getting them to share their problems with the TELEC team and with their fellow teachers, the team is keeping a close contact with the real classroom. By conducting research into user patterns, the team is gaining better knowledge about the teachers so that it can provide them with the kind of support that will help them to develop professionally.

The remaining sections of this paper offer detailed discussion of the design of the three databases:

- test band database (Section 2)
- study ideas database (Section 3)
- grammar database (Section 4)

2. Constructing an Open-Access Testing Database for Teachers

David Coniam

2.1 Preamble

The testing TeleNex database has the objective of supplying teachers with a variety of reliable tests at various levels across the secondary school age and ability range. Work has been conducted in certain other countries to build up test banks for teachers (Van Thiel and Zwarts 1986). The majority of such test database construction has, however, centred around item response theory (IRT) techniques in order to build calibrated item banks than to construct an open-access testing database for teachers. See, e.g. Millman and Arter 1984.

Most teachers in Hong Kong engage in some form or other of testing, in that they produce tests for mid- or end-of-term examination purposes. They then discard them once the test is over, only to have subsequently to produce further tests for the same age-group or similar ability levels.

One of the aims of the TeleNex testing database is therefore to reduce this wastage of effort, to recycle as far as possible (since most schools work to a common syllabus,) so that teachers can draw appropriate and reliable tests off the network when they need them. The question of quality is, however, one that is not to be taken lightly since the fact that a teacher produces a test does not necessarily imply that it will be a good test. This matter, however, is not one that can be addressed in the current paper.

2.2 One Step Forward, Two Steps Back: Calibrated Item Scale

A test produced by a teacher for a particular class, which is then to be fed back into the system, needs to be inserted into the database at a point which is appropriate for the ability levels of other schools and classes. It therefore appeared necessary to take one step backward and examine the merit of constructing a common scale before the above could be achieved. The fact that the intention was to construct a scale of ability suggested that the most appropriate model would be one based on IRT techniques, (see e.g. Wright and Stone 1979, Henning 1984, 1987). The measurement unit in this model is derived from logarithms and is known as a logit: one logit corresponds roughly to one standard deviation.

2.3 The Common Scale

Using tests for Secondary 1 - Secondary 7 with common linking items throughout, a common scale was established. The data was then analyzed using the one-parameter IRT model with the *Bigsteps* computer program (Linacre and Wright 1993). Logit values were re-scaled to a normative mean of 60, with a spacing factor of 9.1, in order to present measurement values that teachers might more easily recognise for comparative purposes. The results are presented in table 1 below.

Table 1

Test Analyses: Rescaled Mean Logit Values of Common Items

S7	88.8
S6	84.9
S5	78.7
S4	70.7
S3	68.4
S2	67.2
S1	60.0

The scale that has emerged shows a range of 28.8 points - 3.16 logits, confirming data presented by Henning (1984) where he describes a rather shorter scale produced from the results of US college students. It can also be seen that the scale is not linear. There is generally half a standard deviation (half a logit) of difference between the forms. This is, however, not the case with the between S2 and S3, where ability differences appears to be minimal. Since the present paper is an operational one about the TeleNex testing database, the reader is referred for further details to Coniam (forthcoming).

2.4 Two Sides to the System

The testing database can therefore be viewed from two angles: the tests teachers can access freely from the database, and the calibrated items they take in order to determine where a test they have produced should fit into the database, or where their students fit into the ability scale, and hence what material is appropriate for them.

For example, a teacher has an S4 class she wants an appropriate test for. She is initially given 10 calibrated items with an approximate mean of 72 (cf. table 1 above) and she then selects testing material from this approximate level. Her students take the test and she sends their answer scripts to the TeleNex testing database team. Upon analysis, her students emerge with a mean of 79.5. We inform her of this fact, suggesting that in future she should select material from around level 79. This, as can be seen is more in the region of S5.

A similar case exists when a teacher produces a test, let us say for an S2 class. She asks the computer for some calibrated items and is given 20 with an approximate mean of 68. Students scripts are sent to us as before. An analysis of this test reveals a mean of 61.1, suggesting that the test the teacher has produced is more fitting an S1 class than an S2, and is therefore inserted into the database at level 61.

In addition to the teacher-produced tests, the Hong Kong Examinations Authority has very kindly consented to letting TeleNex make its past papers available on the database. This is a very useful facility for teachers from two perspectives. It means that not only will past papers now be available to teachers, rather than - as is too often the case - inconveniently disappearing, but also teachers will be able to construct their own mock public examinations by mixing and matching test types out of a number of past papers, rather than being constrained to a single fixed order.

2.5 Selecting/Designing a Test

In order to select a test, a teacher has to go through a number of steps. Assuming that she has taken testing material off the database from a particular class before, the teacher is first presented with the ability level of the class she intends to design a test for: this can be left to the default ability score stored in the database, or the teacher can request material of a more or less demanding nature.

She then selects the material she wants according to the following criteria:

- test item type
- number of items
- topic
- maturity level (While a very able S2 class may be of an approximate S5 level in terms of ability, such a class will not necessarily be sufficiently mature to handle S5 level material. This point needs to be carefully considered.)

The teacher has been making the above selections from menus. It is possible to briefly preview material selected before the final product is assembled.

Once the teacher is happy with all her selections, the final process is to pass whatever material the teacher has selected to the computer's word processor. The whole test is then sequentially renumbered (and questions randomised in the case of multiple-choice questions) and formatted, the name of the school and class are added to the head of the test, and an answer key is generated. The teacher can then re-format the test if she so wishes, adding any special rubric, instructions, etc., before she prints the test out.

3. Considerations and Constraints in Developing a Teaching Ideas Database

Sima Sengupta

This section will describe the 'teaching ideas' database that is being developed as a part of the *TeleNex* network. First a general description of the salient features of the *TeleNex* network and the teaching ideas database will be provided. Then the possibilities and problems of authoring a hypertext database will be discussed. This will be followed by a demonstration of how these factors have influenced the organisation of the teaching ideas database developed so far.

3.1 An Overview: *TeleNex*

This database, as part of an interactive network (*TeleNex*), supports participatory design (Greenbaum and Kyng 1991) where users are able to participate and collaborate with the development team in the process of design. Participation at different levels loosely matches the steps of Good's model (1992) of ideal participatory design consisting of five phases:

- relationship building
- contextual inquiry
- brainstorming
- story boarding
- iterative design

At the relationship building stage some teachers were familiarised with the development process as the design progressed, and these teachers were brought in to comment on the content, lay-out and the organisation. For contextual inquiry and brainstorming, Good's second and third steps were built into a needs assessment study carried out through a questionnaire survey of all the participating teachers. However, the next two steps of the model, story boarding and iterative design, were unnecessary because of three special features of 'participation' offered by *TeleNex*. Firstly, the network allows users to comment on and annotate each screen they read and the host records the details such as the screen annotated, the user, the site and the date. The database team then act on the comments. In addition, the 'teaching ideas' come with an evaluation icon enabling teachers to make overall comments. Thirdly, the system can chart the pattern of use or the path an individual user has followed, i.e. the time the user has spent on each screen, the number of times he/she has gone back to the screen and the way he/she has navigated through the screens. Therefore, this participation intertwines developmental and evaluative activities, where the ultimate shape will be determined by user demands and requirements.

3.2 An Overview: Teaching Ideas

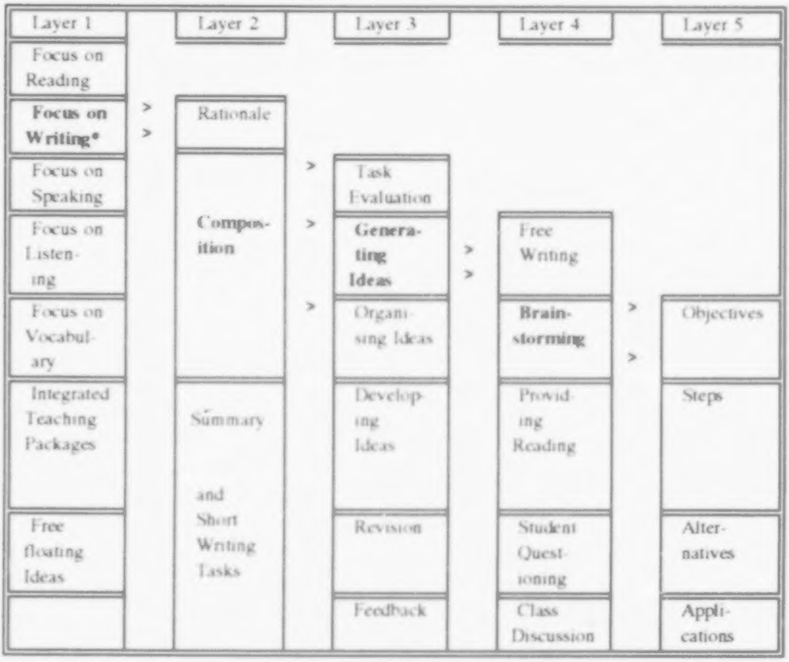
The teaching ideas database for TELEC aimed at providing practical activities which would serve both immediate and long-term needs of the classroom teacher. The first layer of 'teaching ideas' consists of the following 'Overview.'

- Focus on Writing
- Focus on Speaking
- Focus on Listening
- Focus on Vocabulary
- Integrated Teaching Packages
- Free floating Activities

Under these layers further layers are subsumed as pull down menus. An example of the layering within the domain 'Focus on Writing' is presented in figure 2 below. As we can see within 'Focus on Writing', there are four layers of pull down menus and a fifth layer of possible selections. The topmost layer is **Focus on Writing**, which is subdivided into three parts at the second layer. Each item on this layer subsumes another layer, such as the six items under **Composition**. This third layer contains another layer of information, as the example **Generating Ideas** shows. The fourth layer of items such as **Brainstorming** subsumes a final layer consisting of individual screens, which can be selected from a menu, for example by clicking **Brainstorming**. At the time of writing this paper the domain 'Focus on Writing' is still under development.

Figure 2

Focus on Writing: Layers



* Bold letters signify the item clicked by the user and next layer that appears on the screen.

BEST COPY AVAILABLE

3.3 Possibilities and Problems of Hypertext Writing

The endeavour of creating a database has enormous possibilities as well as a few constraints. The writing of a database shares some common grounds with authoring a book; both, at the very least, have to be coherent, adequately researched, well planned and well thought out. However, the similarity ends there.

An author can have an adequate mental representation of a book since the presentation of information is linear. Therefore, the reader can be expected to look up required information by referring to a page or a chapter. The print medium also gives the reader more time to process the data and to move to new or related information.

Hypertext, on the other hand, needs different organisational devices, needs more sensitivity to reader needs and can make fewer assumptions about the reader/user (Denton and Kelly 1992). The hypertext writer is not bound by linearity and this makes a variety of organisational alternatives possible. However, other kinds of constraints are posed by attempting to map the hierarchy of information, the depth of layers, the nodes or chunks of information that should be grouped together and the links that have to be established between these chunks for easy navigation (Woodhead 1991). Consideration of these constraints have played a role in the organisation and writing of this database although from the very outset the focus has been and still is on exploring how to make adequate use of technology to support the conceptual content of the database (Pugh 1993).

3.4 Practical Implications: Organisation

The possibilities offered by non-linear presentation of information has meant that 'Focus on Writing' can incorporate multiple organisational perspectives. Users have the following alternative entry points:

- They can enter at the level of 'Overview', which introduces them to the overview of the database. Layer 1 in figure 2 can be the starting point that users click for browsing through the layers. For example, a teacher wishing to get some ideas about teaching writing will choose **Focus on Writing**.
- They can choose to get into the database through the kind of writing they want to teach, i.e. through layer 2 in figure 2, and then make their selection.
- They can also enter the database to find teaching ideas for a particular dimension of writing, for example how to teach organisational strategies for composition (layer 3 in figure 2).
- They can also choose a particular activity as an entry point, such as **Brainstorming** for generating ideas.

Other entries will be possible through the search mechanism:

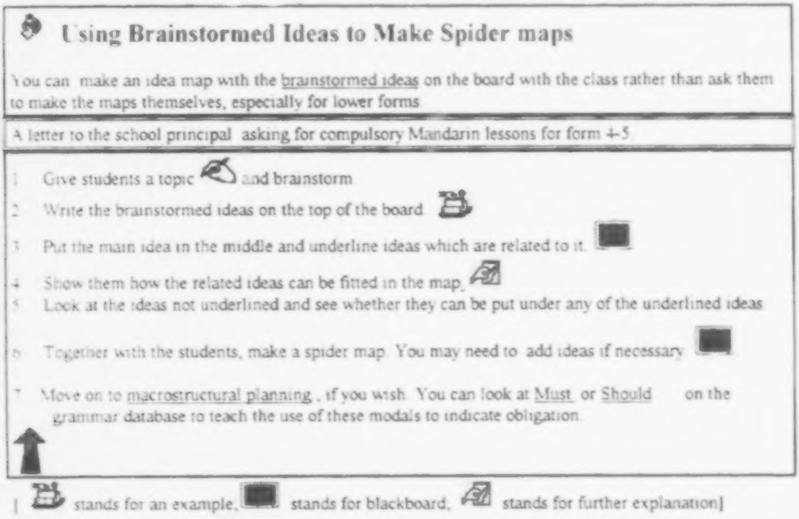
- Users are also able to make topic-based choices, i.e. they can look at the variety of writing topics, such as a descriptive topic, and see how such a topic can be presented to students. This is layer 5 (**Alternatives**) in figure 2.

- Users may also choose to enter by selecting either a particular textbook they are using or an item from a past examination paper, and follow it through.

There are two types of links. Firstly, there are pop-up screens which explain a point. Icons are used to indicate these pop-up screens. Secondly, there are jumps, indicated by underlines (on the screen they are underlined in green) which take the user to another part of the database. For example, users reading a screen about how to teach organisational strategies for composing can jump to the 'Pedagogic Grammar' database to see how teaching of a grammatical point can fit in with the teaching activity suggested on the specific screen the user is reading. The following example screen from **Organising Ideas: Spider Maps** shows some of these features. The users can jump to **Generating Ideas** by clicking **Brainstormed Ideas** at the beginning, or to the grammar database by clicking **Must** at the end, or to other ways of teaching planning or elaborating by clicking their choices. The icon in 1, when clicked, will display the topic as shown on the screen below.

Figure 3

A Sample Screen in the Teaching Ideas Database



Using Brainstormed Ideas to Make Spider maps

You can make an idea map with the brainstormed ideas on the board with the class rather than ask them to make the maps themselves, especially for lower forms.

A letter to the school principal asking for compulsory Mandarin lessons for form 4-5

1 Give students a topic and brainstorm

2 Write the brainstormed ideas on the top of the board

3 Put the main idea in the middle and underline ideas which are related to it.

4 Show them how the related ideas can be fitted in the map.

5 Look at the ideas not underlined and see whether they can be put under any of the underlined ideas

6 Together with the students, make a spider map. You may need to add ideas if necessary

7 Move on to macrostructural planning, if you wish. You can look at Must or Should on the grammar database to teach the use of these modals to indicate obligation.

[stands for an example, stands for blackboard, stands for further explanation]

3.5 The Way Ahead

This section has described the first exploratory steps taken in an ongoing odyssey through tools such as hypermedia, and futuristic concepts such as building up a participatory teacher support through interactive computer-mediated communication. Carlson (1988) points out that hypertext is an electronic means for enhancing four primary categories of idea processing: reading, annotating, collaborating and learning. All these categories have a role to play in the *TeleNex* network. The future is full of rich possibilities and potential.

4. Designing a Hypertext Pedagogic Grammar of English

Wu Kam-yin

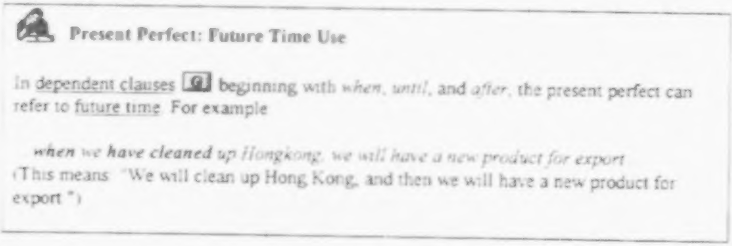
As a resource for participating teachers, a *Hypertext Pedagogic Grammar of English* (henceforth *Pedagogic Grammar*) is being developed for *TeleNex* as the grammar database. In this section, I shall describe some of the features of this grammar, and examine how its intended readership has affected the form it takes. But first, a definition of ***hypertext grammar*** is in order.

4.1 Hypertext Grammar

To explain what a hypertext grammar is, let me define ***hypertext*** first. Hypertext is a computer term, and refers to a method of organizing texts. More specifically, hypertext "stores text in a non-linear structure of nodes and links" so that it is possible to think of "information existing in a multi-dimensioned space" (Davis 1993:18). In contrast, conventional magazines, newspapers, novels and so on organize texts in a linear structure.

A ***hypertext grammar***, as its name indicates, is a grammar written in a hypertext format. The following example, from *Hypertext Pedagogic Grammar of English*, gives a flavour of how it works. The screens have been slightly modified for this paper.

Suppose a teacher is reading the following information on a computer screen:

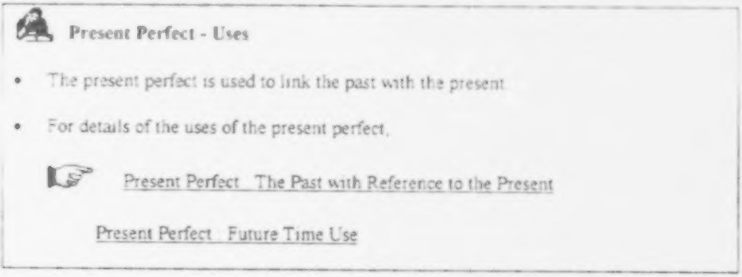


Present Perfect: Future Time Use

In dependent clauses beginning with *when*, *until*, and *after*, the present perfect can refer to future time. For example

when *we* ***have cleaned*** *up Hongkong, we will have a new product for export*
(This means "We will clean up Hong Kong, and then we will have a new product for export.")

When she comes across any underlined items on which she needs more information, she can, by pressing a button on a mouse, access the required information immediately. In other words, from the screen above, the teacher can go go the following two screens:

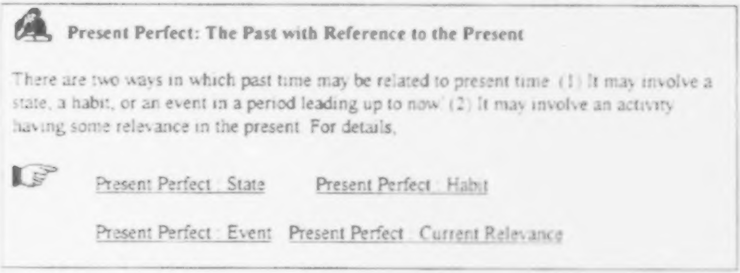


Present Perfect - Uses

- The present perfect is used to link the past with the present
- For details of the uses of the present perfect,

☞ Present Perfect : The Past with Reference to the Present

Present Perfect : Future Time Use



Present Perfect: The Past with Reference to the Present

There are two ways in which past time may be related to present time (1) It may involve a state, a habit, or an event in a period leading up to now' (2) It may involve an activity having some relevance in the present For details,

☞ Present Perfect : State   Present Perfect : Habit

Present Perfect : Event   Present Perfect : Current Relevance

Notice that each of these two screens is linked to other chunks of information, which the teacher can easily retrieve by, again, simply pressing a button on her mouse.

Having explained what a hypertext grammar is, I will now discuss three advantages that it has over conventional reference grammars:

1. The users of a hypertext grammar have more control over what to read; they do not have to follow any pre-set, linear structure of texts found in conventional reference grammars.
2. Information in a hypertext grammar, presented on stand-alone computer screens, is more easily digestible.
3. Retrieval of information is faster and easier, and this facilitates the looking up of cross-referenced materials. In this connection, it is worth noting that grammatical terms are mutually defining and thus in any grammars, extensive cross-referencing is essential (Greenbaum 1987:192). An efficient means of retrieving cross-referred items makes a hypertext grammar more user-friendly.

The *Hypertext Pedagogic Grammar of English* that is being developed for *TeleNex* has all these advantages and more. It is not simply a grammar in a computerized form. It is a grammar specially written for secondary schools English language teachers in Hong Kong, and contains information that we believe will be useful to them. I describe some of its other major features below.

4.2 Features of *Hypertext Pedagogic Grammar of English*

Terminology

We define carefully all the grammatical terms appearing in *Hypertext Pedagogic Grammar of English* for two reasons. First, we do not assume any technical knowledge about English linguistics on the part of the users. Second, there is a lack of standardization in the meaning of grammatical terms, which are sometimes used to mean different things by different grammarians. Thus, it is essential that all grammatical terms be defined clearly and unambiguously.

How much new terminology should be introduced is a difficult problem to solve. In Hong Kong, a lot of teachers are familiar with traditional grammatical terms, but a number of these terms are either misleading or unsatisfactory. For example, "present perfect" is less transparent in meaning than "past in present", a term used by Halliday (1985), but teachers are more familiar with the former term than with the latter. For this reason, we retain the term "present perfect" and explain that its meaning is past in present.

In other cases, unfamiliar terms such as "theme", "rheme", "reference" and so on are useful in talking about English texts, and we do not hesitate to use them. All these terms are defined carefully in *Hypertext Pedagogic Grammar of English*.

Authentic Examples

The need to use authentic texts in analyzing English has been properly emphasized and clearly explained by researchers such as Sinclair (1987) and Willis (1990). At the Teachers of English Language Education Centre, a corpus of authentic English has been built up. This consists of appropriately 1 million words of newspaper English and another 1.2 million words of spoken English. *Hypertext Pedagogic Grammar of English* draws its examples from this corpus, and this has enabled us to see, among other things, that the present perfect can be used to refer to future time, a meaning that is not discussed in many reference grammars.

In drawing examples from the corpus, *Hypertext Pedagogic Grammar of English* does not cite only isolated, decontextualized sentences. With the help of technology, the grammar stores in separate computer files the original texts from which example sentences have been taken. This allows the readers, if they so wish, to examine the use of authentic language in authentic contexts.

Focus on Meaning

Greenbaum (1988:45) observes that the intended readership of a grammar affects the form it takes. *Hypertext Pedagogic Grammar of English* is intended for secondary schools English language teachers in Hong Kong. In general, most of these teachers are familiar with linguistic form, but not as many of them know that grammar is, in the words of Halliday (1985:xvii), "a system for making meaning." Thus, in developing a grammar for these teachers, we pay more attention to function rather than structure. As an example, we explain only briefly how the various tenses in English are formed, but describe in detail how they can be exploited to express meanings. Using authentic texts, we demonstrate how more than one tense can co-occur in the same piece of writing. The advice sometimes given to students, that they should not shift from one tense to another in a single text, is simply a misconception.

4.3 Teaching Implications

The intended readership of *Hypertext Pedagogic Grammar of English* has influenced the form it takes in another way. We think that teachers consulting this grammar will be interested not only in understanding how the English language works, but also in finding out what teaching implications can be drawn from the grammatical description provided. For this reason, we discuss teaching implications where appropriate. These implications are based on a number of considerations, e.g. how frequently a grammatical item is used in English texts; how important it is to enable students to see the regularities in English grammar; and so on. To illustrate, while not all grammarians agree that the various uses of the simple past tense express remoteness (see e.g. Palmer 1974), we suggest that teachers teach this global meaning to students. Our rationale is that this will help students to realise that the different uses of the simple past are not unrelated and that there are regularities in grammar.

4.4 Teaching Ideas

Knowing what teaching implications to draw from grammatical analyses is one thing. Knowing how to present grammatical information to students is another. In *Hypertext Pedagogic Grammar of English*, we provide a bank of activities to enable teachers to teach grammar communicatively. A lot of these activities, supplied by teachers, have been tried out in Hong Kong classrooms and students have reacted favourably to them. We have good reasons to believe that they will work in other classrooms too.

Once the database is fully operational, we shall revise the grammar in the light of feedback from teacher users, so that it will be as useful to English teachers as possible.

Acknowledgement

This is an annual survey conducted by the Education Department of Hong Kong. The survey referred to was conducted in 1991. The authors wish to thank the Education Department for granting Dr Amy Tsui permission to use the raw data from the Department's 1991 survey of teachers, quoted in the first part of this paper. The figures quoted there are the results of the author's analysis of the raw data, in a research project funded by the Research Grants Council on the academic and professional background of ESL teachers in schools in Hong Kong.

References

Carlson, P.A. (1988). Hypertext: a way of incorporating user feedback in to on-line documentation. In E. Barrett (ed.), *Text, context and hypertext writing with and for the computer*. Cambridge, MA: MIT Press.

Coniam, D. (forthcoming). Designing an ability scale for English across the range of Hong Kong secondary school forms. *Hong Kong Papers in Linguistics and Language Teaching*, 17.

Davis, S.B. (1993). Hypertext and multimedia. *English Today*, 33, 9(1):17-24.

Davis, N. (1991). The lack of electronic communications in initial teacher training. *Journal of Computer Assisted Learning*, 7:2-6.

Denton, L. and Kelly, J. (1992). *Designing, writing and producing computer documentation*. New York: McGraw Hill.

Good, M. (1992). Participatory design of a portable torque-feedback device. In L. Borman & B. Curtis (eds.), *Human factors in computing systems: CHI '92 conference proceedings*. Amsterdam: North-Holland.

Greenbaum, S. (1987). Reference grammars and pedagogical grammars. *World Englishes*, 6(3):191-197.

Greenbaum, S. (1988). *Good English and the grammarian*. London: Longman.

Greenbaum, J., and Kyng M. (1991). *Design at work: cooperative design of computer systems*. Hillsdale, NJ: Erlbaum.

Halliday, M.A.K. (1985). *An Introduction to functional grammar*. London: Edward Arnold.

Henning, G. (1984). Advantages of latent trait measurement in language testing. *Language Testing*, 1:123-131.

Henning, G. (1987). *A guide to language testing*. Cambridge, MA: Newbury House.

Ip, W.C. (1989). A study of the potential of electronic bulletin boards as perceived by teachers. M.Ed. dissertation. Faculty of Education, The University of Hong Kong.

Linacre, M., & Wright, B.D. (1993). *Bigsteps, Version 2.4*. Chicago: Mesa Press.

Mangan, K. (1992). Computer networks help public schools forge new ties with higher education. *The Chronicle of Higher Education*, September 1992.

Meadows, J. (1992). International collaborations in teacher education: a constructivist approach to using electronic mail for communication in partnership with schools. *Journal of Information Technology for Teacher Education*, 1(1):113-125.

Millman, J. & Arter, J.A. (1984). Issues in item banking. *Journal of Educational Measurement*, 21:315-330.

Palmer, F.R. (1974). *The English Verb*. London: Longman.

Pugh S.L., (1993). Using case studies and collaborative computer-assisted communication to support conceptual learning in a teacher education course on critical reading. *Educational Technology, 33(11).*

Sinclair, J.M. (ed.) (1987). *Looking Up* (pp. 150-159). London: Collins.

Tsui, A.B.M. (1993). TELEC Proposal, *Language Campaign Bulletin*, 7:6-7.

Van Thiel, C.C., & Zwarts, M.A. (1986). Development of a testing service system. *Applied Psychological Measurement*, 10:391-403.

Willis, D. (1990). *Lexical syllabus*. London: Collins.

Wright, B.D., & Stone, M.H. (1979). *Best test design*. Chicago: Mesa Press.

PART FIVE

COGNITIVE PROCESSES IN LANGUAGE LEARNING

INTRODUCTION

Interest in the processes of language learning was particularly strong at the conference, as illustrated by a large number of contributions in this area. The papers selected for inclusion are all based on experimental studies. The majority of the papers involve learners whose L1 is Chinese and the various studies are concerned with L1 and L2 learning and teaching, involving both English and Chinese. An interesting feature of the collection is the extent to which the processes involved in learning English and Chinese are compared within the same study. The papers cover three broad topics:

1. strategies used by second language learners,
2. language learning processes,
3. teachers' awareness of students' language learning.

Two papers by *Peter Gu* and *Chiou-lan Chern* discuss the strategies used by language learners. A number of contributions at the conference were based entirely on the results of questionnaires given to learners to report of their own language learning strategies. These have generally not been included in this collection, since such studies are already well represented in the learner strategies literature. Gu's paper involves an in-depth study of the L2 (English) vocabulary-learning strategies of two Chinese learners, both university students; one demonstrated considerable expertise as an L2 learner, whereas the other was less successful. The subjects' performance on reading tasks which involved dealing with unfamiliar vocabulary was investigated by means of think-aloud protocols. The analysis of the data focuses on the way the subjects react to unfamiliar words, their patterns of dictionary use and the reinforcement strategies used to commit new words to long-term memory. *Chern* also based her paper on a study comparing the strategies of two types of learners. The paper reports on a study into Chinese readers' metacognitive awareness in reading Chinese and English. The study involved the subjects in three tasks: a pre-reading interview about their own reading behaviour, a reading task and a retrospective interview based on their experience of the reading task.

The impact of cultural schemata on reading comprehension and recall is investigated in *Terry Dolan's* paper, which reports the results of an experiment which compares British and Hong Kong Chinese students' reading performance on an English text, presented under different conditions, including the use of illustrations and schemata which reflect both British and Hong Kong cultural backgrounds. The study suggests that Chinese readers are more likely than British readers to be influenced by illustrations which accompany texts. Based on the Hong Kong subjects' results, *Dolan* argues for the use of both English and Chinese illustrations and schemata in the presentation of English texts to Hong Kong students.

Yilin Sun's paper is concerned with word recognition in Chinese, both as L1 and L2. The paper reports an experiment which examined the effects of word frequency and stroke complexity upon the processes of word recognition. *Sun* is interested in establishing whether the features which influence word recognition in Chinese as a foreign language are the same as those identified in previous L1 studies. With the notable exception of *Paul Meara's* work in UK, word recognition research has largely been restricted to L1.

Monica Hill also looks at vocabulary and addresses, in particular, the issue of phonological processing in word learning. Her paper reports the results of an experiment which involved presenting new vocabulary to learners through different conditions. Phonological processing is represented in the experiment both by direct hearing of new vocabulary and through the decoding of phonemic symbols. Her results support the view that long-term retention of vocabulary is assisted if phonological processing takes place during the initial encounter with new words and suggest that not only oral but also written recall are enhanced by paying attention to the phonological dimension of new words.

Tasks which help to mediate between texts and learners' responses to them are the topic of the paper by *Desmond Allison, Vivien Berry and Jo Lewkowicz.* Their paper reports on an experiment which investigated the effects of different mediating tasks upon students' writing of a summary of a reading text. The authors are interested, in particular, in establishing whether students' understanding of a reading text is improved if they discuss the text before undertaking any formal response to it. The influence of follow-up discussion upon comprehension is compared with the more traditional practice of asking students to answer written comprehension questions based on the text.

The influence of learners' personalities upon their performance in oral tests is addressed by *Vivien Berry.* Her paper reports the results of two studies which investigated the ways in which different combinations of personality types in pair and group interviews effect individual performances. *Berry's* data suggest that different combinations of personalities can produce different results and that particular test tasks may favour (or disadvantage) particular personality types. In view of the increasing use of the group oral test, particularly in public examinations, *Berry* emphasises the need for further research into the role of personality variables in language testing.

Finally, three papers look at ESL teachers' awareness of language and language learning. *Stephen Andrews* investigates teachers' knowledge and perception of grammar and their attitudes towards it. His paper, which is based on questionnaire data from 101 native speakers of Cantonese and 29 native speakers of English and involves comparisons with the results of similar studies conducted in UK, sets out to uncover what teachers actually understand by grammar. *Andrews* speculates that the confused state of grammar teaching in Hong Kong schools might be a reflection of the differing perceptions of grammar among members of the teaching profession and advocates that need for follow-up studies on this topic.

Teachers' ability to identify sources of lexical difficulty for their students in pedagogical texts is the topic of *Arthur McNeill's* paper. *McNeill* compares the performance of native and non-native speaker ESL teachers on a lexical decision-making task and concludes that teachers who speak their students' L1 are generally better at predicting the difficulties their students will face when reading particular texts. The ability of native English speakers to tune in to ESL students' actual vocabulary difficulties appears to vary widely, even among experienced teachers with advanced qualifications in education and applied linguistics. *McNeill* suggests that a possible explanation for the poor decision-making by this group might be some wrong assumptions about word difficulty derived from the literature on reading and vocabulary.

Margaret Falvey's paper also discusses teachers' ability to identify sources of difficulty for language learners. *Falvey* reports on a study which required trainee teachers to give concurrent verbal reports while reading a text. By means of the data derived from the verbal reports, *Falvey* aims to provide trainee teachers with greater insights into the processes involved in reading and thus to become sensitive to the demands of reading tasks upon language learners.

Arthur McNeill

VOCABULARY LEARNING STRATEGIES OF GOOD AND POOR CHINESE EFL LEARNERS[1]

Gu Yongqi

Ask any foreign language learner about his headaches in learning the language and one thing you will surely get is the difficulty in remembering words. And yet it is these words that make up a language, and ample evidence suggests that the vocabulary size of a learner is highly predictive of his entire language ability (e.g. Gui 1985, Meara and Jones 1987). The development of vocabulary in a foreign language is undeniably one of the most crucial and yet difficult issues that researchers can ill afford to overlook.

Since Meara's (1980) call for vocabulary research in applied linguistics, the last decade has seen a rapid development along this line. However, with the exception of work on contextual guessing, few empirical studies have, as yet, targeted the vocabulary learning mechanism. What we have on vocabulary learning processes are more often than not prescriptive in nature. For example, the linguistic paradigm on vocabulary learning tends to draw inferences from the descriptions of the target language at various levels. Richards' (1976) excellent analysis of what it means to know a word, Seibert (1945) and Clarke and Nation's (1980) detailed accounts of the linguistic and logical structures that underlie contextual guessing procedures, are all illuminating examples in providing the necessary thinking on what needs to be learned. Nevertheless, as Van Parreren and Schouten-Van Parreren (1981) rightly observe, these linguistic analyses are not psychological processes as such and hence do not tell us how vocabulary is learned. While there are some perceptive descriptions of the learner's target language vocabulary (e.g. McNeill 1994, Meara and Ingle 1986, Palmberg 1987), these studies give us insights on the output of the learner's interlanguage lexicon. We still do not know the mechanisms by which these lexicons came into being. And yet, it may well be true that the ways a learner learns vocabulary determine the retrievability and flexibility of his lexicon and, to a considerable extent, his overall language achievement (see Ahmed 1989). In other words, research is badly needed on the strategies and processes of vocabulary development that very possibly make the good learners good and the poor learners poor.

Research on memory mnemonics in remembering foreign language words is mostly done in the fields of experimental and educational psychology. Not surprisingly, psychologists' interest is almost entirely on memory strategies per se. To them, foreign language words are not much different from nonsense words or any meaningless materials for the purpose of experimental learning. While some psychologists (e.g. Beck et al. 1987) realize the complexity of 'the complete vocabulary knowledge', others (McDaniel and Pressley 1989) still maintain that 'regardless of the objectives of a vocabulary program, a key ingredient is often acquisition of vocabulary-word-definition associations' (p. 204). By far the most extensively tested vocabulary learning strategy is the keyword method devised by

Atkinson (1975), and results all point to a single conclusion, that the keyword method is superior to almost all other types of strategy (see Cohen 1987 and Paivio and Desrochers 1981 for comprehensive reviews). Nonetheless, the fact that the keyword mnemonic is still confined to laboratory truth after twenty years of rigorous research speaks for itself as to the applicability of mnemonics in the foreign language classroom. There is one single question the experimental psychologists cannot and perhaps would not care to answer, 'Has anybody ever successfully learned a foreign language using the keyword method or any other mnemonic devices?'

Unfortunately, few linguists who work on language learning strategies have bothered to focus on vocabulary. And perhaps more seriously, research on language learning strategies has been far too quantitatively oriented. Granted the reliability of questionnaire research that so characteristically epitomizes a large proportion of empirical studies in this area,[2] it takes more than strategy counts and correlations to understand why some learners gain by painstaking efforts whereas others fail despite laborious attempts. As has been repeatedly demonstrated (Gu 1992), using more varieties of strategies and using them more frequently may not necessarily guarantee success in language learning. How one uses a strategy may be just as important, or even more important, to learning than the number of strategies one employs. It is thus the contention of the present author that qualitative methods may start to reveal exactly where quantitative methods fall short. For example, both the successful and the unsuccessful learners will look up unfamiliar words in the dictionary, and they may well report similar frequencies of dictionary use in answer to a questionnaire item. However, they may differ dramatically in terms of which word to look up, when to look it up, how to look it up, what to look up, and what to take down after they have looked up a word. As Ahmed (1989) demonstrated, it may well be these micro strategies and processes of learning that determine the success or failure of the learning outcome.

With these contentions in mind, the following study is designed to explore and describe the authentic strategies and processes of vocabulary learning employed by Chinese learners of English and to see in qualitative terms if these processes are in any way related to the result of learning. In so doing, three stages are examined, i.e. how a new vocabulary item is handled during the initial encounter; how it is then looked up in the dictionary, if at all; and how it is reinforced afterwards.

Method

Subjects

This article reports part of a larger study that based its sample on 24 Chinese learners drawn from a total population of 978 third-year non-English majors learning English at Beijing Normal University. By the time the study took place, these learners had all had six years' experience of learning English as a course in their secondary schools around China, plus more than two years of English learning experience at BNU. Expressed in hourly terms, this would mean that a typical third-year non-English major would have spent more than 1,212 classroom hours (932 in secondary schools, and 280 in university) learning English as a foreign language (The State Education Commission 1986a, 1986b).

Two learners (one 'good', the other 'poor') were chosen for a case study in this report. They were selected on the grounds that first of all, they had similar backgrounds before entering university, both coming from key secondary schools in urban environments, and secondly both reported high levels of motivation and desire to learn English and had spent numerous extracurricular hours on English learning, and yet while Learner One was highly successful (getting a score of 96.5 in the national College English Test, Band 4), Learner Two suffered miserably with a score of only 31 in the same test. Intelligence-wise, although Learner 2 did not perform as well as Learner 1 who boasted a score beyond the 95 percentile point among young urban Chinese of his age on Raven's Standard Progressive Matrices (SPM), she was fairly competent for nearly reaching the 75 percentile point on the SPM. It was thus felt that their striking difference in English achievement could have resulted from the ways they had been learning English, including, of course, the ways they had been learning vocabulary.

Instruments

Two reading passages[3] (see Appendix 1) were selected for intensive reading,[4] Text 1 for the top group, and Text 2 for the bottom group.[5] A pilot new word density analysis among 13 randomly selected third-year non-English majors at the same university revealed that Text 1 had a familiar to unfamiliar word ratio of 43.7:1, while Text 2 had a ratio of 91.3:1, indicating roughly that for typical students of the same grade, Text 1 contained about twice as many new words as Text 2.

The texts were broken up into. meaningful segments marked with small red strokes which acted as reminders for the subjects to stop reading and verbalize their thinking processes. Think aloud protocols on the reading processes as well as on vocabulary learning while and after reading were obtained from both subjects. In addition, immediate retrospective interviews based on the researcher's field notes were also conducted immediately after each task was finished so as to capture anything of interest that could not be revealed by the think-aloud process.

Procedures

The study as reported here was completed in one session. Each subject began with a think-aloud training of roughly 30 minutes using another text that was rated in the previously mentioned pilot study as easier than Text 1 but more difficult than Text 2. They were then told to read the text the way they would normally do when preparing for the Intensive Reading lesson and to verbalize anything that was going through their head, even when they needed to look up a word in the dictionary and when they took down notes. The researcher prompted them along the way with 'What are you thinking now?' or 'Could you tell me how you arrived at that conclusion?' when the subjects fell into silence and when they failed to verbalize a point. The subjects were then asked a few questions pertinent to their respective passages concerning things they failed to verbalize. Next, they were requested to do what they normally do with vocabulary items that had been identified as unfamiliar. They were provided with enough pieces of paper for note taking and for whatever other purposes pertinent to vocabulary learning. This stage was also followed by retrospective questions wherever the need arose. The whole process of vocabulary learning starting from identifying a problem word in the reading passage to the reinforcement of the word after reading was tape recorded with the subjects' permission.

Analyses

The think-aloud recordings were next transcribed for analysis, with a keen focus on vocabulary learning processes and strategies. The three stages of vocabulary learning that were built into the original design served as the framework for analysis. Namely, the initial handling of a problem word after its identification, the dictionary strategies the learner revealed when checking the word, and the reinforcement strategies the learner used in order to commit the word to his/her long-term memory. The transcripts were then studied meticulously to see how each subject went about learning vocabulary and to uncover any strategies and processes within each of the above-mentioned stages that could distinguish the good learner from the poor one. A descriptive model in the form of a flow chart was then drawn up to illustrate each learner's processes of vocabulary learning. The two learners were finally compared and contrasted stage by stage on the metacognitive as well as cognitive levels.

Findings

Learner 1: Male, 21

Vocabulary learning through reading.

Learner 1 read his passage three times. Firstly, he went through the passage trying to get the gist of what was being said and underlined words that were unfamiliar to him, words he thought he would go back to later. He guessed at the meaning of each of his unfamiliar words at this stage and did not bother to stop reading and check them up in the dictionary. Secondly, he glanced through the passage and scanned for his underlined words or any other unfamiliar words he had overlooked during the first reading. He then looked up these words in his dictionary and located the meaning that he thought was appropriate to the context. For words that were important and interesting to him, he would look for their usages, and other meanings and usages that had little to do with the context. Occasionally, he would also browse through the same page in the dictionary where the target word lay, just to see if there were any other words that were of particular interest to him, e.g. words that resembled the target word in spelling or sound and were easily confused with the target word. And in extreme cases, he would become so interested in a totally irrelevant word that he went on to look that word up. He took two types of notes along the way. For words that he thought were especially interesting and useful to him, he would note down the meanings, usages, and sometimes examples on a piece of paper he had been instructed to use as his notebook. Also included in his notes were pronunciations of words that were thought to be difficult (e.g. sewerage) and synonyms from the text and from his own vocabulary repertoire (e.g. hem: n. edge, fringe; v. hem in: enclose, surround*). For the rest of the unfamiliar words that were useful for text comprehension, he would write their meanings in the margins or between the lines of the original text. Finally, after completing all these procedures, he would go through the whole text again very quickly, focusing on an overall understanding of the passage, pausing only when special attention was needed on certain words or phrases that were thought to be worthy of emphasis.

Vocabulary learning after reading.

Next, Learner 1 demonstrated what he would normally do to reinforce vocabulary items he identified as unfamiliar during reading. Thirty five vocabulary items in all (including words and phrases, totalling 8.43% of the entire text) had been identified as either unfamiliar or partially familiar. By far, his emphasis was on the words and phrases he noted down in his notebook (the separate sheet of paper). First of all, he would take a quick glimpse of his definition/explanation of each item either in Chinese or in English and raise his head to recall the original English word/phrase. For words that were long and difficult to spell, he would also scribble them down rapidly on a piece of paper. In addition, he attempted to recall everything he had gone through to understand a particular item, from the contextual meaning to other related or unrelated meanings, from words that looked similar, synonyms, to phrases and examples he had found out in the dictionary. He even made up sentences of his own using some of the items that were of special interest to him. Finally, he went through his list very swiftly two more times, going firstly top-down and then coming back bottom-up. The whole process, interview time excluded, took him roughly 90 minutes. The following is an example showing all the procedures Learner 1 went through in learning the verb 'smart'.

Example 1: Smart
Text Segment No. 03 (Text No. 01)

Protocol	Strategies
[guesswork]	
S: Smog SMARTS *the eyes and chokes the senses*, SMART, I know it's an adj, but here it must be a verb, a verb, *SMART the eyes* must be hurting the eyes. I'm of the world not absolutely sure about it, though, should have it confirmed later, needs to be carefully studied when I have time later. [A015-017]	Inferring from part of speech and knowledge of the world Postponing
[dictionary work]	
S: SMART is usually an adjective, it doesn't seem an adjective here. SMARTS, (reads in dictionary). SMART, SMART is locate meaning definitely not an adj here, so I'll go for the verb. Oh, there's such a meaning for SMART, it means to sting. No?	Using part of speech to locate meaning
SMARTS the eyes, oh, yes, it is to sting, it's this meaning then. SMART, let me see if there're any set expressions that go with it. Ah, I see an adj with a similar meaning here. 'a SMART blow, a SMART blow', a good beating [wrong in dictionary], this, need to remember this.	Negotiation of meaning Extended dictionary use
R: Why do you think you need to remember this?	
S: This 'SMART blow' is, you know, somehow I feel it's commonly used, so I gave it a look. Actually I didn't do it with much intention. When I saw this word, this giving sb a good blow, seems to be a commonly used expression, for instance, I gave him	Selective attention (personal interest)
a smart blow, I should remember how it is said [in English].	Monitoring
'a SMART blow' means a good blow, I need to go back to the verb, for verb, it's only the meaning to sting. Oh, let me see the example sentences.	Looking for usage
Wha what's this? [?] it can be an intransitive verb, and then, feeling painful, [?] 'with, from, from [?]', here's an expression, I don't want to remember it, can't remember everything anyway.	Selective attention
Are there any other expressions? 'as SMART as a new pin', very handsome [wrong in dictionary], ahh, this I'll remember. Very handsome is something quite often used. 'Pin' seems interesting to me, like a needle. 'as SMART as a new pin', 'pin' does seem interesting, like a needle.	Personal interest

I'll look it up. 'Pin' I know it's sort of a needle, as handsome as a needle? it doesn't seem, (reads in dictionary) 'pin', here it is, 'pin', needle, panel pin, [?], tiny things, pin through, sting, limit, accuse, [?], no, I can't solve my problem. Then I've got no other choices but to remember it [as it is], 'cause a pin is nothing but a pin. I'll remember it then. Let me see if there're any others. [?] 'as SMART as a new pin', eh, nothing else, I'll remember this then. (writes down under smart **as ~ as a new pin**) Right, then, so much for this, now I'll come back,	Extended dictionary use
'SMARTS the eyes', apparent by now, it's stings the eyes painfully. [A194-237]	Monitoring

[Reinforcement 1: during 3rd reading]

S: Smog SMARTS the eyes and chokes the senses, now here, when I meet SMART, when I read SMARTS the eyes, I tell myself to remember it, to	
remember SMARTS the eyes, it's to sting the eyes painfully. And also CHOKE the senses, 'cause I remember I took it down in the notes, now I better reinforce it. It's no more than telling myself to pay attention to it, and I'll certainly read on. [B097-101]	Reinforcement while reading

[Reinforcement 2: after reading]

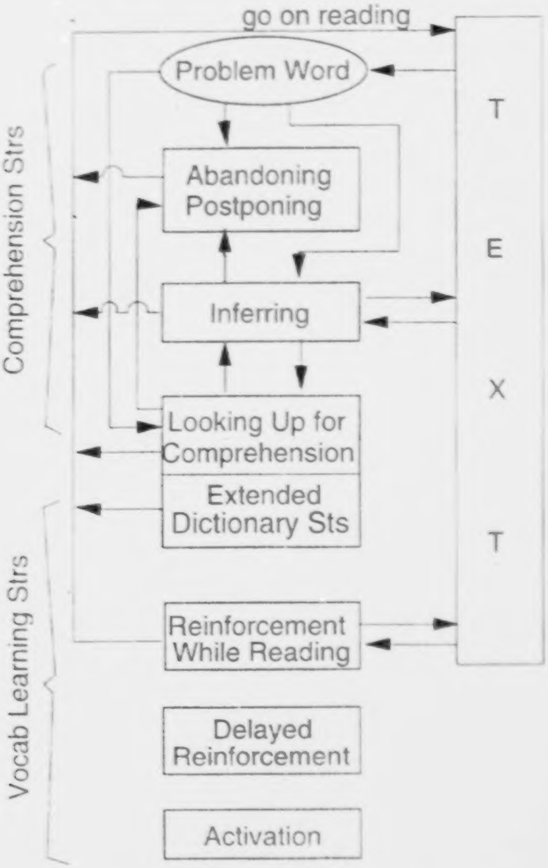
S: (reads the Chinese equivalents while recalling the	
original English words) [see immediate retrospection for confirmation] SMART, I have it here as a good blow, 'a SMART blow', now, let me try to use it. "give somebody a SMART blow. I gave him a SMART blow yesterday" (laughs). I	Delayed reinforcement
gave him a good beating yesterday, useful word (laughs). Sting, sting, now I remember, SMART the eyes, it's what's in the text. This is easy, 'give	Activation
sb a SMART blow', I think I've remembered it. Stings the eyes, it's also easy. Oh, there's another,	Evaluation
eh, 'as SMART as a new pin', eh, very handsome, interesting, very handsome, "He, He is as SMART as a new pin", this is interesting. "He is", he's very handsome, "He is as SMART as a new pin." So much for this. [B347-361]	Activation

A descriptive model.

Based on the previous analysis, the following descriptive model is drawn up to illustrate more clearly and dynamically the process of vocabulary learning Learner 1 went through (See Figure 1). Most of the procedures in Figure 1 can be borne out by previous descriptions and the example above, others are traceable in the protocols.

Figure 1

Vocabulary Learning Through Intensive Reading: Descriptive Model of a Good Learner



When Learner 1 identified a problem word in the text, he would immediately abandon it if it was a word of no significance either to text comprehension or to his personal interest (e.g. the Greek Premier's name). He could also postpone the understanding of a word to a later stage. For words the meaning of which he thought impossible to infer, he would go to the dictionary. In most cases, however, he used various clues to guess the meaning of a word from its context. This strategy sufficed for a very general understanding of the text, and satisfied his first aim of reading, i.e. to answer the question 'Why is the city dying?' Nonetheless, he was not completely contented with his understanding of the passage until he had checked those words he was still not sure of in the dictionary. In Figure 1, I have labelled his strategies so far as Text Comprehension Strategies.

What followed epitomizes how 'reading to learn' is possible. Learner 1 did not stop after making sense of the text, he went on to consult his dictionary to find out how an interesting word could be used, how this word was related to other words he knew. And moreover, he would consciously tell himself to remember a word when it appeared again later in the same text. In other words, Learner 1 was not just reading, he was deliberately learning the words he regarded as meaningful and useful to him.

After making sure he had understood both the passage and the crucial vocabulary words, Learner 1 did reinforcement as well, rehearsing his notes several times and relating new words to his existing vocabulary stock. He also found it interesting to make up his own sentences using some of his favourite new words. To distinguish these strategies from the Text Comprehension Strategies mentioned above, I have dubbed them Vocabulary Learning Strategies. In reality, however, vocabulary learning starts right from the outset when a problem word is identified as important or interesting.

Learner 2: Female, 22

Word for word translation as reading.

Learner 2 read her passage only once, and stopped at every unfamiliar word to check its meaning in her dictionary. To be more exact, quite a proportion of her unfamiliar words were in fact familiar to her, e.g. ability, create, rather, certain, no longer, etc. These were usually words she had met and rehearsed time and again. On no occasion, however, did Learner 2 demonstrate any guessing strategies to infer any of the unknown or partially known words. In using her dictionary, Learner 2 rarely had any problem locating a word, but she had serious problems locating the right meaning, especially under multiple-meaning entries. Her strategy at this point was to 'find the general or common meaning' under that particular entry (Tape Position: 01B050-238) and write it down along the margins or between the lines. And that was about everything she did for note-taking. No part of speech, no pronunciation, no usage, no examples, nothing but the Chinese equivalent, albeit not always the right one. After making sure that every unfamiliar item within a sentence had been looked up, she would then try to understand the sentence by stringing together all these words, now their Chinese 'equivalents', along the order of the original English sentence. When this did not make any sense, as was very often the case unless the English syntax of the sentence under question was exactly the same as its Chinese translation, her strategy, though subconsciously applied, was to impose Chinese syntax upon the string of Chinese words now still in the order of English syntax, and reconstruct the often meaningless word-chain into a more or less meaningful sentence by adding or deleting any words as were necessary and occasionally shifting the order of words as needed. This done, she would feel she had already understood the sentence, disregarding the fact that she still could not string the sentences thus made into a meaningful text. She did attempt, though in vain, on a number of occasions, to postpone the understanding of the current sentence until she finished reading the following sentence(s) and/or until she read the preceding sentence again. After about two hours of painstaking effort (interview time excluded), however, no single sentence had been completely understood. The following is an example showing how Learner 2 decoded a sentence.

Example 2
From Learner 2
Text Segment No. 20 Tape Position: 2A318-2B018

Problem word identified: Visual
Other possible problem items: what is more, hard, sense, appearance

S:	*What is more, it becomes hard to make sense of one's own,* don't know the word that follows. V-I-S-U-A-L, (looks it up), seems to be sight [wrong part of speech]? (silence). I don't know *what is more,* how can I put it, what is a lot? what is more than a lot? Isn't it too [?] (laughs), somehow I feel it should be more complicated than this direct translation of mine. What is it that is more then? It	Looking for meaning only, disregarding part of speech 'What-is-more': Word-for-word translation
	becomes, *become hard to,* (long silence), *make sense of one's,* use your own sight, *make sense,* (thumbs through dictionary) [end of use of? side A of Tape 2]	Mixed up with 'make use of'?
	More is, use the, make your own sight, and the sensation is made? Oh, it's to make the, work hard to make the, sort of sight, one's own sight. This is what I feel. How do I say it?	Changing word order Adding and deleting words
R:	Do you feel you've understood it?	
S:	I feel I have.	
R:	One's visual what?	
S:	One's visual sight becomes, eh, pretty difficult.	

To illustrate more clearly the process by which Learner 2 went about making sense of an English sentence, the previous example is crystallized into the following three steps:

Steps Used by Learner Two to Decode a Sentence

Step 1: Translate every translatable word, leave out the function words such as 'to' and 'of' as well as words that do not seem important. In short, get the meaning of each word, and do not bother about anything else.

Step 2: String these meanings together and see if they make sense, if not, make up a meaningful sentence by either changing the word order where necessary, or by adding or omitting words when needed.

Step 3: Continue refining the sentence using the strategies in step 2 until it becomes a reasonably acceptable, at least syntactically acceptable, Chinese sentence.

These procedures first resulted in something that bore much resemblance to some unsophisticated machine translation, which, after the second and third steps, turned into Chinese sentences more or less of the 'Colourless green ideas sleep furiously' type.

Whatever the analogy, perhaps all previous illustrations can be boiled down to the following implicit guideline this learner appeared to be following. Each word has a definite meaning which comes either from the word list at the end of each unit in a textbook or from the dictionary, and that things seem so fluid and arbitrary to her beyond the word level that she has to manipulate word orders in order to make the resulting sentence a more meaningful one. In effect, it is no overstatement to say that Learner 2 was imposing meaning onto text rather than extracting meaning out of text. Why can this be possible, and after more than eight years of English learning? The way she remembered vocabulary provides further clues.

Mechanical rehearsal as vocabulary learning

When asked to show what she would normally do with the 43 vocabulary items (8.88% of the entire text) she identified as problematic, Learner 2 did exactly what she often did with vocabulary appearing in her textbooks. Firstly, she wrote down each problem word two to seven times with its Chinese equivalent copied beside it one to five times. These words were arranged in the order they appeared in the text, e.g. the word 'locate' appearing far away from 'location'. Though she did murmur each item she was copying, letter by letter, followed by pronouncing the whole word and its Chinese translation, she recalled later that the murmuring itself did not serve much purpose, it was the shape and spelling of each word together with the Chinese version that were what she was trying to internalize.

After finishing every item in this manner, Learner 2 did another 6 rehearsals. The second rehearsal began when she copied down all the items she had just rehearsed in one column on the far left and their Chinese equivalents beside them in another column. Next, she covered the English version with a piece of paper and tried to recall it by looking at the Chinese, taking off the piece of paper from time to time for items she could not recall. After that, she covered the Chinese and attempted to recall it from the English. And this process continued until she did the seventh rehearsal. Somehow she arbitrarily stopped copying the bottom half from the third rehearsal on, perhaps realizing she was wasting too much time. The entire process of rehearsing lasted about 90 minutes.

A descriptive model

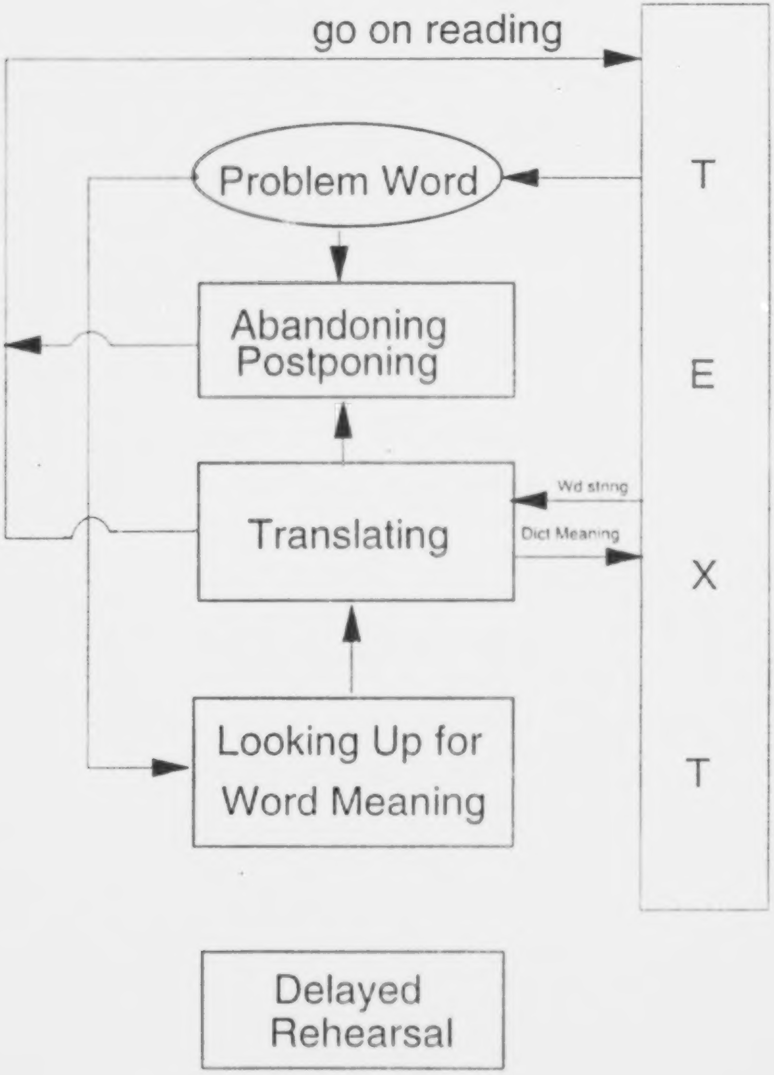
Learner 2's whole process of intensive reading and vocabulary learning is summarized in figure 2.

When a problem word was identified in the text, Learner 2 would either ignore it or go to the dictionary immediately, and she went to the dictionary for nothing but lexical meaning. As a result, the definition she got from the dictionary was often not the meaning appropriate to the context. And yet, instead of negotiating between the dictionary and the text for a suitable definition, Learner 2 would impose onto the text a 'general meaning' she derived from the dictionary. When the resulting string of Chinese 'equivalents' did not make sense, she would impose Chinese syntax onto it and, as it were, forcibly add or omit words in order to make it sound like a sentence. This completed, she would have a false judgment that she had already understood the sentence. Even when she did not believe in her own interpretation, she would either abandon the sentence or 'wait until the teacher explains it later.' 'And if the teacher doesn't, it's not important anyway.'

No intentional reinforcement was done during reading, although she so often expressed deep concern over the words she had rehearsed so many times and still had to resort to dictionary. Reinforcement only came after reading, and it was only in the form of mechanical rehearsals. No attempt was made to relate these words to words she already knew, to the context where it appeared, or to any syntactic roles that the word could play.

Figure 2

Intensive Reading and Vocabulary Learning: Descriptive Model of a Poor Learner



Moreover, for Learner 2, learning English had almost been tantamount to remembering word lists. In a subsequent interview, she revealed that she viewed vocabulary as the most important part of English learning, and she had spent over sixty percent of all her English learning time on reciting the word lists at the end of each lesson exactly the way she did for me. 'I just write them like that, again and again, on any piece of paper I can get hold of.' Worse still, she had been spending most of all her university study time on English! 'I can't remember what else I have done for the past two years and more, whenever I go to the study room, the only things I carry would be my English textbooks.'

Discussion

Comparing the Two Learners: Where Do the Differences Lie?

Anyone experienced in analyzing expert and novice behaviours will know that 'the good are simply good and the poor simply poor' (R.K. Johnson, and Q.F. Wen, personal communication 1993). They are different anyway one looks at them. I will confine myself, however, to focusing on the metacognitive and cognitive processes in which Learners 1 and 2 demonstrated dramatic differences in their vocabulary learning (See Table 1). In so doing, I will briefly touch upon another important aspect as well, i.e. the emotional state each learner was in and the influence of this upon the learning process.

At the metacognitive level, the expert learner (Learner 1) saw intensive reading as a process of learning as well as information decoding. He was therefore actively aware of any learning opportunity during reading by constantly relating all vocabulary items in front of him to his own lexical stock. He evaluated the familiarity of every item and determined its relative importance and hence the level of processing the item needed (in this case, whether an item needed to be attended to, abandoned or guessed, checked in the dictionary for meaning, studied more carefully for usage, reinforced during or after reading, or even activated for firmer control). All these decisions were made according to two criteria:

1. an item's relevance to text comprehension;
2. its relevance to his personal interest.

The pace and scope of learning was also carefully monitored so that he would not go too far away from completing the central task of reading and learning within a period of time that was reasonable both to his general time management and to his judgment of the experimental condition. In addition, he was very aware of how well he was going through each step, and made decisions as to the amount of extra time and energy he needed to reach his target. He did err from time to time, but overall, he was in comfortable control.

Table 1

Comparing the Two Learners

Strategies	Expert Learner	Novice Learner
Word selection	1. relevance to text comprehension 2. relevance to personal interest	Non-selective
Abandoning/ Postponing	highly selective	rarely at the word level; always at and beyond the sentence level
Contextual Inference	1. frequent 2. using varieties of cues	rarely using cues, no successful attempt, only word for word semantic matching between L1 and L2
Using dictionary to comprehend	negotiating between text and dictionary meaning - fitting dictionary meaning into text	impose dictionary meaning onto text
Using dictionary to learn	yes, and highly selective	No
Reinforcement while reading	yes, and highly selective	No
Delayed Reinforcement	1. encoding of declarative + procedural knowledge of word 2. rehearsal of declarative + procedural knowledge of word	rehearsal only and then only of declarative knowledge of word
Activation	yes, and highly selective	No

The novice learner (Learner 2), on the other hand, did not have a clear idea of what intensive reading was for, and aimed for only 'a rough idea of a passage.' She was thrown into a vicious circle in which she had to spend so much time decoding the passage that she would be left no time learning whatever language points she thought important to learn; and yet, the less time she spent on the usages of these items, the less English syntax she would be able to reconstruct in her mind, and hence the more she would have to cling to Chinese syntax for comprehension, which in turn would result in more time, less comprehension, and therefore more frustration. As a result, in order to understand more, she was desperate to go for every word that came without due evaluation as to its importance to comprehension, let alone her personal interest. Apparently, the poor learner was overwhelmed, not by words she did not know, but by not being able to make sense of words she thought she knew. She was monitoring very little, not because she did not have the ability to control the timing and scope of her learning process, but because she lost control psychologically, and was not even able to use her common sense. In fact, she had to stop for a while when reading the third paragraph, and had to be comforted for a few minutes in order to go on. Likewise, she seemed unable to evaluate her own learning appropriately in much the same way as a desperate drowning person would not be able to evaluate the usefulness of a straw. This is understandably aggravated by the cruel fact that after learning English this way for more than eight years, she would certainly be frightened at any incomprehensibility and would hence easily enter the state of desperation. 'I see this as my last chance', she said in a fainting voice, referring to her participation in my research. In fact, metacognitive control over learning and the attached emotional feeling of being in or out of control can never be overemphasized, so much so that disregarding these two aspects would render any interpretation of the following cognitive strategies void.

At the cognitive level, the good learner selected vocabulary items for different purposes, while the poor learner dealt with the same list of problem words throughout the whole process. For example, while the good learner was highly selective when choosing words to abandon or postpone, the poor learner almost never abandoned a word and looked up every word she thought problematic. However, she had to abandon a sentence or postpone her understanding of it simply because she probably had to do it anyway. In addition, the good learner frequently tried to use a variety of cues in order to guess the meaning of a problem word before looking it up, the poor learner rarely used any cues and made no successful attempt. The dictionary was used by the good learner as an aid to comprehension and a source to learn from. When a word was being looked up, one could see him negotiating between dictionary explanations and contextual meaning. His purpose was to find an appropriate dictionary meaning and fit it into the context. The poor learner, on the other hand, took the dictionary as a collection of nearly absolute lexical meanings and tried to find a 'general meaning' under a dictionary entry and impose it onto the text. And a dictionary to her served only this purpose and nothing more. While the good learner had the leisure to stop at any interesting point while reading and to tell himself which word needed more attention to be reinforced, the poor learner was too much engrossed in puzzling out the general meaning of the text to bother about the learning of any word, not to mention its reinforcement. When it came to remembering vocabulary items identified in the text, the good learner managed to relate new items to relevant familiar items in his repertoire, to the context where the item appeared, and to his knowledge schemata in general. In so doing, his attention was not just on linking a lexical form with its corresponding meaning, he also showed interest in remembering the usages of the word. He rehearsed as well, in much the same manner as the poor learner did, but while he did it with meaning and in addition to encoding strategies, the poor learner did nothing but mechanical rehearsal, without even relating the noun form and the verb form of the same word. And more importantly, the content of the poor learner's rehearsals was almost exclusively the connection between a written symbol and one or two, as it were, fixed explanations, which explains why she confused 'imagine' with 'image', and 'shape' with 'sharp'. It was small wonder then that when the good learner was able finally to activate an item, the poor learner was profoundly baffled for failing to retrieve words from her mental jungle of unrelated and arbitrary associations between written forms and their dictionary meanings.

Theoretical and Practical Implications

Word knowledge: the least a learner needs to acquire. Since Richards (1976), applied linguists have nearly come to a consensus as to what it should mean to know a word. To most theorists (e.g. Richards 1976, Carter 1987, McNeill 1994), knowing a word in a second or foreign language means knowing the form and structure of the word, the semantic, affective, and pragmatic meanings associated with the word, the syntactic behaviours of the word, the likelihood of encountering the word in normal discourse, and how the word is associated with other words in the target language lexicon. To most foreign language learners, however, this is too much to expect of them. For example, the affective, stylistic, and pragmatic entailments of a word are simply luxuries the beginning foreign language learner cannot afford. The full knowledge of every word should be the ideal end toward which a learner ought to strive for, and therefore may not necessarily be needed to succeed in using the target language with considerable ease. Are there any essentials, then, without which a learner will most probably fail?

One of the questions I asked in a subsequent interview was 'What does it mean to you to have learned a word?.' Learner 2 did not hesitate to reply 'I think it's to have remembered it. I can recognize it when I see it,' whereas Learner 1 gave the following thoughtful answer: 'To have learned a word doesn't just mean to know its meaning. It's best to put it in a context, to be able to use it in various contexts, for instance, what sort of a situation or a state the word describes, how it is used, and with what words it collocates.' Obviously, to the poor learner, learning a word means to memorize the form and meaning association, but to the good learner, it means not only seeing the form and meaning association in contextual lights, it also means 'putting the word in a context', i.e. anchoring the form and meaning association in a sentence. Reflecting back on what each learner did to learn vocabulary, I conclude with the following tentative proposition: so far as vocabulary learning is concerned, the least a learner needs to acquire is the form, the referential meaning and the basic syntactic behaviour of each word.

Dynamic vocabulary competence. A vocabulary in a language is never the simple addition of individual words with static meanings listed out in dictionaries (Richards 1976, Carter 1987). It is a whole dynamic network of interrelated words each playing a semantic, syntactic and pragmatic role when activated. This would mean that even with the simplest form of comprehension, the dynamic vocabulary competence has to be activated in order to determine the types of entailments of each word in a sentence.

To date, however, the lion's share of attention on vocabulary learning strategies has been given to the memory of form-meaning pairs, as if the meanings associated with a particular form were static. And moreover, rarely do we see any warning to the learners that committing vocabulary items to memory should not be an end in itself. While I do not oppose the employment of memory strategies in vocabulary learning, I do see the danger of overemphasizing the memory of form-meaning word pairs (cf. Carter 1987), for this would give some learners the false impression that remembering the form and its corresponding native language equivalent is all that is required of vocabulary learning. And in turn, inadequate vocabulary learning strategies might be induced due to the previous false impression. Unfortunately, while poor memory strategies and/or lack of more efficient ones will only affect the pace of vocabulary learning, inadequate understanding of what vocabulary is, and the consequent lapse of attention on the dynamic nature of words, which involves their contextual entailments and syntactic behaviours, would, as has been shown in Learner 2 in this study, result in very serious consequences. For example, the good learner's vocabulary, however small and however incomplete in the various aspects of vocabulary knowledge discussed earlier, is one that is dynamic and alive, but the poor learner can hardly be said to have an English vocabulary, as her repertoire of words are mostly unrelated forcible connections between orthographic forms and one or two of their dictionary definitions.

How can the poor be helped?

Elsewhere (Gu 1992) I have likened the language learning strategy researcher's efforts to Robin Hood's mission of robbing the rich to feed the poor. This study, however, has thrown this endeavour into doubt. If it is only the number of strategies and the frequencies of strategy use that deprive the poor of improvement, providing them with more alternatives obtained from the strategically rich and telling them to use their strategies more often would easily solve the problem. The real picture is unfortunately much more complicated. The poor learner in this study employed a much narrower range of strategies, but she probably also used the rehearsal strategy, for instance, more often than most people did, and hence had become so used to it that she would fall back on it after trying other strategies that had no immediate effect. Will it be much help telling her to negotiate between dictionary explanations and the contextual meaning when in fact she cannot make sense of the context? Similarly, showing her how the good learner went all the way to looking up words irrelevant to the text he was reading would probably do more harm than good.

Poor learners like Learner 2 have usually been paralysed by their inadequate strategies for too long. For people who have never been on their own feet before, telling them to walk this or that way certainly would not help. Learner 2 needs to remember not just the meaning of a word, but to relate the word with other words she knows, and above all she should pay attention to the usage of words. These and other good strategies can only work if she learns how to choose the right words to focus on, and how to monitor and evaluate her own strategy use as well as her learning process. In other words, she needs a systemic treatment that takes everything analyzed in this article into account, i.e. at least in EFL learning, she needs special education care. To help learners like her, researchers and especially teachers would be better off taking up the role of a nurse, pushing the wheel chairs of the strategically disabled, guiding them, encouraging them patiently to stand up again.

Specifically, to give her English a touch of life, Learner 2 needs to be helped to realize that words are dynamic in nature, and that learning a foreign language is far more than remembering the target language equivalents of all native language words. She should therefore pay special attention to the syntactic behaviours of each word (cf. Nattinger 1980, 1988). As a first step, she must form the habit of knowing at least the part of speech of every word she tries to remember. She then needs to read as much as she can and intentionally cultivate a sense of, as it were, English grammar in action, as opposed to the rules she learned in grammar books. Finally, after she can make sense of English sentences without translation, she should be encouraged to develop contextual guessing strategies and pave the way for vocabulary acquisition through reading.

To help enhance her memory, she should be guided to organize her repertoire in both paradigmatic and syntagmatic dimensions (Meara 1984) so that vocabulary learning is diverted from simple addition of new items and mechanical rehearsal of existing ones to an integration of old and new items by means of reconstruction, i.e. she should shift from surface level to a deeper level of processing (Craik and Lockhart 1972). Finally, she should change from her mainly visual memory of words to a combination of visual and acoustic memory, so as to release the burden of letter strings on short-term memory, and to add a helpful method of encoding (cf. Hill 1993).

Conclusion

This article scrutinizes in detail the vocabulary learning processes of two Chinese EFL learners. It shows how the way students deal with vocabulary can, to a considerable extent, contribute to success or failure in foreign language learning.

It was demonstrated that in order to succeed, one must take the dynamic nature of words into consideration so that the resulting interlanguage is made up of a vocabulary that is alive. By contrast, seeing vocabulary as nothing but paired associations between form and definition is a sure way to fail.

It should be noted that while I would not claim the representativeness of these two learners, I do not wish to play down the seriousness of some of the reported problems either. Given the enormous population of English learners in China, even if Learner 2 represented only a tiny proportion (though my intuition suggests otherwise), it would still include tens of thousands of people. Therefore, besides providing estimates of these and other vocabulary related problems, future research should logically focus on finding possible solutions, be they providing the poor with what we get from the rich, or helping the poor develop their own self-sufficient vocabulary learning strategies.

Notes

1. This article is based on part of the data taken from the author's Ph.D. project. The author wishes to thank Dr. R. Keith Johnson for his patient encouragement and enlightening supervision. Special thanks are also due to Mr. Arthur McNeill, Prof. John Biggs, and so many others at the Faculty of Education of HKU whose friendship and guidance have made my work possible and enjoyable. I am also indebted to Prof. Robert B. Kaplan who generously offered invaluable comments to an earlier draft of this paper. Finally, thanks also to all subjects at Beijing Normal University whose kind cooperation provided me with insightful data.

2. But see Gu, Wen, and Wu (1993) for a detailed discussion on reference ambiguities and other problems of the Likert-scale in research on learning behaviours.

3. Both passages were taken from Walter's (1982) *Authentic Reading*. In fact, Text 1 was adapted from an article in *Time* magazine, and was about pollution in Athens; whereas Text 2 was an article introducing the felt image people have about their bodies and was adapted from an article in *The Guardian*.

4. Intensive reading in China means much more than reading intensively. Its aims are at least twofold: reading to comprehend and reading to learn, with the latter being probably the most important form of English learning. In fact, a course entitled *Intensive Reading* often provides the major source of English input.

5. Paradoxically, we lose comparability at the textual and linguistic level by using two texts instead of one; on the other hand, we reduce the comparability at the strategy level if only one text is used. Two texts were chosen in this study in view of the present focus on vocabulary strategies. Readers interested in this issue are referred to Kletzien (1991) for further arguments.

6. In this example, 'hem', 'hem in', and 'fringe' appeared in Text 1 and were identified by Learner 1 as new vocabulary items, whereas 'edge', 'enclose', and 'surround' were words he retrieved from his own lexical repertoire. Notice he was not only regrouping and hence recoding words semantically for himself, linking new words with words he already knew, he was also classifying words according to their grammatical functions, here, their part of speech.

7. See Appendix 2 for the legends of transcription and translation.

References

Ahmed, M.O. (1989). Vocabulary learning strategies. In P. Meara (ed.), *British studies in applied linguistics 4: Beyond words* pp. 3-14. BAAL/CILT.

Atkinson, R.C. (1975). Mnemotechnics in second-language learning. *American Psychologist*, 821-828.

Beck, I.L., McKeown, M.G. & Omanson, R.C. (1987). The effects and uses of diverse vocabulary instructional techniques. In M.G. McKeown & M.E. Curtis (eds.), *The nature of vocabulary instruction* pp. 147-163. Hillsdale, NJ: Erlbaum.

Carter, R. (1987). *Vocabulary: Applied linguistic perspectives*. London: Allen & Unwin.

Clarke, D.F. & Nation, I.S.P. (1980). Guessing the meaning of words from context: Strategy and techniques. *System* 8:211-220.

Cohen, A.D. (1987). The use of verbal and imagery mnemonics in second-language vocabulary learning. *SSLA* 9:43-62.

Craik, F.I.M. & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behaviour* 11:671-684.

Gu, Y. (1992). Robin Hood in SLA: What has the learning strategy researcher taught us. Unpublished manuscript, University of Hong Kong.

Gu, Y., Wen, Q., & Wu, D. (1993). How often is often: Reference ambiguities of the Likert-scale in language learning strategies research. Paper presented in July 1993 at the Guilin ELT International Conference, Guilin.

Gui, Shichun (1985). Woguo yingyu zhuanye xuesheng yingyu cihuiliang de diaocha he fenxi [A survey analysis of the English majors' vocabulary size in China]. Xiandai Waiyu [Modern Foreign Languages], 27(1):1-6.

Hill, M. (1993). A word in your ear: To what extent does hearing a new word help learners to remember it? Paper presented at the International Language in Education Conference in December 1993. Hong Kong.

Kletzien, S.B. (1991). Strategy use by good and poor comprehenders reading expository text of different levels. *Reading Research Quarterly*, 26:67-86.

McDaniel, M.A. & Pressley, M. (1989). Keyword and context instruction of new vocabulary meanings: Effects on text comprehension and memory. *Journal of Educational Psychology*, 81:204-213.

McNeill, A. (1994). Vocabulary testing: Defining word knowledge. In P. Falvey & J. Boyle (eds.) Language testing in Hong Kong. Hong Kong: Chinese University of Hong Kong Press.

Meara, P. (1980). Vocabulary acquisition: A neglected area of language learning. *Language Teaching & Linguistics: Abstracts*, 13:221-246.

Meara, P. (1984). The study of lexis in interlanguage. In A. Davies, A. Howatt, & C. Criper (eds.), *Interlanguage*, pp. 225-235, Edinburgh: Edinburgh University Press.

Meara, P. & Ingle, S. (1986). The formal representation of words in an L2 speaker's lexicon. *Second Language Research*, 2:160-171.

Meara, P. & Jones, G. (1987). Tests of vocabulary size in English as a foreign language. *Polyglot* 8(1):1-40.

Nattinger, J.R. (1980). A lexical phrase grammar for ESL. *TESOL Quarterly*, 14:337-344.

Nattinger, J.R. (1988). Some current trends in vocabulary teaching. In R. Carter & M. McCarthy (eds.), *Vocabulary and language teaching*, pp. 62-82, London: Longman.

Paivio, A., & Desrochers, A. (1981). Mnemonic techniques in second-language learning. *Journal of Educational Psychology*, 73:780-795.

Palmberg, R. (1987). Patterns of vocabulary development in foreign language learners. *SSLA*, 9:201-220.

Richards, J.C. (1976). The role of vocabulary teaching. *TESOL Quarterly*, 10:77-89.

Seibert, L.C. (1945). A study of the practice of guessing word meanings from a context. *Modern Language Journal*, 29:296-323.

The State Education Commission, PRC. (1986a). *English teaching syllabus for full-day secondary schools.* The People's Republic of China: People's Education Press.

The State Education Commission, PRC. (1986b). *English teaching syllabus for arts and science students in tertiary institutions.* Shanghai Foreign Languages Education Press.

Van Parreren, C.F., & Schouten-Van Parreren, M.C. (1981). Contextual guessing: A trainable reader strategy. *System,* 9:235-241.

Walter, C. (1982). *Authentic reading.* Cambridge: Cambridge University Press.

Appendix 1

The Two Texts in Segmented Form

Text 1

01 Stinking buses, their passengers pale and tired, jam the crowded streets.
02 Drivers shout at one another and honk their horns.
03 Smog smarts the eyes and chokes the senses.
04 The scene is Athens at rush hour.
05 The city of Plato and Pericles is in a sorry state of affairs, built without a plan, lacking even adequate sewerage facilities, hemmed in by mountains and the sea, its 135 square miles crammed with 3.7 million people.
06 Even Athens' ruins are in ruin: sulphur dioxide eats away at the marble of the Parthenon and other treasures on the Acropolis:
07 As Greek Premier Constantine Karamanlis has said,
08 'The only solution for Athens would be to demolish half of it and start all over again.'
09 So great has been the population flow toward the city that entire hinterland villages stand vacant or nearly so.
10 About 120,000 people from outlying provinces move to Athens every year,
11 with the result that 40% of Greece's citizenry are now packed into the capital.
12 The migrants come for the few available jobs, which are usually no better than the ones they fled.
13 At the current rate of migration, Athens by the year 2000 will have a population of 6.5 million, more than half the nation.
14 Aside from overcrowding and poor public transport, the biggest problems confronting Athenians are noise and pollution.
15 A government study concluded that Athens was the noisiest city in the world.
16 Smog is almost at killing levels:
17 180-300 mg of sulphur dioxide per cubic meter of air, or up to four times the level that the World Health Organization considers safe.
18 Nearly half the pollution comes from cars.
19 Despite high prices for vehicles and fuel ($2.95 per gallon), nearly 100,000 automobiles are sold in Greece each year;
20 3,000 driver's licenses are issued in Athens monthly.
21 After decades of neglect, Athens is at last getting some attention.
22 In March a committee of representatives from all major public service ministries met to discuss a plan to unclog the city, make it livable and clean up its environment.
23 A save-Athens ministry, which will soon begin functioning, will propose heavy taxes to discourage in-migration,
24 a minimum of $5 billion in public spending for Athens alone, and other projects for the countryside to encourage residents to stay put.
25 A master plan that will move many government offices to the city's fringes is already in the works.
26 Meanwhile, more Greeks keep moving into Athens.

BEST COPY AVAILABLE

27 With few parks and precious few oxygen-producing plants, the city and its citizens are literally suffocating.

(415 words)

Text 2

1 When you close your eyes and try to think of the shape of your body, what you imagine (or, rather, what you feel) is quite different from what you see when you open your eyes and look in the mirror.
2 The image you feel is much vaguer than the one you see.
3 And if you lie still, it is quite hard to imagine yourself as having any particular size or shape.
4 When you move, when you feel the weight of your arms and legs and the natural resistance of the objects around you,
5 the 'felt' image of yourself starts to become clearer.
6 It is almost as if it were created by your own actions and the sensations they cause.
7 The image you create for yourself has rather strange proportions:
8 certain parts feel much larger than they look.
9 If you poke your tongue into a hole in one of your teeth, it feels enormous;
10 you are often surprised by how small it looks when you inspect it in the mirror.
11 But although the 'felt' image may not have the shape you see in the mirror, it is much more important.
12 It is the image through which you recognize your physical existence in the world.
13 In spite of its strange proportions, it is all one piece,
14 and since it has a consistent right and left and top and bottom,
15 it allows you to locate new sensations when they occur.
16 It allows you to find your nose in the dark, scratch itches and point to a pain.
17 If the felt image is damaged for any reason
18 -- if it is cut in half or lost, as it often is after certain strokes which wipe out recognition of one entire side --
19 these tasks become almost impossible.
20 What is more, it becomes hard to make sense of one's own visual appearance.
21 If one half of the felt image is wiped out or injured, the patient stops recognizing the affected part of his body.
22 It is hard for him to find the location of sensations on that side,
23 and, although he feels the doctor's touch, he locates it as being on the undamaged side.
24 He loses his ability to accept the affected side as part of his body, even when he can see it.
25 If you throw him a pair of gloves and ask him to put them on, he will glove one hand and leave the other bare.
26 And yet he had to use the left hand in order to glove the right.

27 The fact that he can see the ungloved hand doesn't seem to help him, and there is no reason why it should.
28 He can no longer reconcile what he sees with what he feels
29 -- the ungloved object lying on the left may look like a hand,
30 but, since there is no felt image corresponding to it, why should he claim the object as his?

(484 words)

Appendix 2.

Transcription Legends

Original text in English	*Italics*
Word in Focus	CAPITALIZED
Subject Talk (direct translation from Chinese)	S: normal
Researcher Talk (direct translation from Chinese)	R: normal
Subject Write	**bold**
Researcher Comment	[ ]
Researcher Observation	( )
Recording Unclear	[?]
Irrelevant	[...]
Omission	(...)
Inserted English Words	' '
Subject Emphasis	Underline
Made-up Sentences	" "

CHINESE READERS' METACOGNITIVE AWARENESS IN READING CHINESE AND ENGLISH

Chiou-lan Chern

Introduction

Metacognition in reading refers to readers' background knowledge of the text, their awareness of using strategies and of the importance of particular strategies. Three metacognitive processes related to reading have been identified by McNeill (1987): self-knowledge, task-knowledge, and self-monitoring. Self-knowledge refers to learners' own perceptions and feelings about themselves as readers that affect their performance; task knowledge refers to the understanding of when to use what strategies; and self-monitoring refers to the awareness of a comprehension breakdown and the knowledge of what to do about it.

Several recent studies have been designed to see if reading comprehension can be improved by increasing learners' awareness and use of strategies in reading. Studies in first language acquisition have proved that increased metacognitive awareness can lead to better use of reading strategies by children (Paris, Cross and Lipson 1984). Researchers in second language education (Barnett 1988, Devine 1984, Kern 1988, Padron, Knight and Waxman 1986) also pointed out that proficient ESL readers showed more awareness of their use of strategies in reading English than less proficient ESL readers. Carrell (1989) also found relationships between readers' metacognitive awareness of various reading strategies and their reading ability in both first and second languages.

Interviews have been commonly used to tap readers' metacognitive knowledge of their own reading processes. Pre-reading interviews can generate readers' reports of their reading behaviors in general; whereas post-reading interviews elicit retrospection of a specific reading task. This study used both pre-reading and post-reading interviews to test the hypothesis that more proficient language learners are more aware of their strategies used in reading. The design of the study is discussed in the following section.

Research design

Participants

Twenty-eight native speakers of Mandarin Chinese who had learned English as a second language were recruited from a university in Taiwan. Participants in this study represented two ESL reading levels. Experienced ESL readers were recruited from the fourth year university students majoring in English; the Inexperienced ESL readers were recruited from the first year university students studying in 11 different departments.

Materials

Two manipulated passages, one in Chinese and one in English, were used as reading materials. The manipulated English material was a passage of 238 words about birds' instinctive ability to fly. Eight function words were taken out of the original passage and replaced by pseudo words which retained the morphological features of English and met the syntactic requirements in the sentence. The manipulated Chinese passage has 486 characters (298 words). Eight of the original words were replaced by nonsense words. Five of these nonsense Chinese words were compounds composed of two real characters which formed meaningless combinations; the other three nonsense words each contained a pseudo character formed with one semantic and one phonetic component in a left to right symmetry.

Procedure

Before reading the two passages, all participants were interviewed individually in their native language, Mandarin, about their general reading habits in reading Chinese and English. The twelve questions used in the interview were adopted from Goodman, Watson, and Burke (1987) and are included in Appendix A. After reading the passages, each participant was also asked to report retrospectively on what they did while encountering difficulties in reading the two passages. Both the pre-reading and post-reading interview sessions were audiotaped for later transcription and coding.

Coding of Data

Participants' responses to pre-reading interview questions were coded by the author. The coding form (See Appendix B), adapted from that of Bruinsma (1990), was translated into Chinese to facilitate coding. The author first determined where multiple responses to one question occurred. All responses were entered and coded except when a participant initially gave a vague response and later specified a strategy, then the first vague response was disregarded and the more specific one was coded.

A second coder who did not know the participants was asked to code all the interview questions. This second coder, a native speaker of Chinese with an MA degree in TESL, was familiarized with the coding form and used the same distinction for multiple responses. The discrepancies in coding were mutually resolved between the coders.

Participants' responses to post-reading interviews were categorized into five categories (See Appendix C) and coded following the same procedures as identified for pre-reading interviews.

Results of the study

A. Pre-reading Interviews

The results of this interview are reported below. The percentages of each category of response by Experienced and Inexperienced groups to the 11 questions are listed in Table 1.

As seen in Table 1, when asked "What is reading?", three-quarters of the Inexperienced ESL readers indicated that reading was meaning-oriented (i.e. a cognitive act). Experienced ESL readers considered reading as either a meaning making process (41.7%) or a classroom/object-related procedure (41.7%). As to the purpose of reading (Question 2), Experienced readers considered reading as either for functional purposes (91.70%) or for enjoyment (50%). The majority of Inexperienced readers considered reading as mainly functional (81.3%) and their response to reading as enjoyment was lukewarm (12.5%). The responses to Question 3 "What do you think makes a good reader?" were more congruent for the two groups of readers: they both considered procedural aspects of reading (speed/accuracy) and meaning/memory as important qualities for good readers.

When asked to evaluate their own reading proficiency in general, most participants rated themselves as "good" readers but none of them considered themselves as "very good" readers. In response to Question 5, "What would you like to do or could you do that would make you a better reader?", both groups of readers regarded practice as the best means to improve comprehension.

The different reading behaviors in reading an L1, Chinese in this case, and an L2, English, as reported by readers themselves can be seen from answers to questions 6 to 10. When reading in Chinese and encountering a new word (Question 6), both Experienced and Inexperienced ESL readers would usually ask somebody or consult a dictionary. Quite a few readers would ignore the unfamiliar Chinese word (43.8% for Inexperienced and 25% for Experienced ESL readers). About one-quarter of the Experienced and Inexperienced readers used context to guess meaning. None of the Experienced ESL readers would bother marking or writing down the difficult Chinese words.

When encountering a new word in reading English, both groups of readers again would ask somebody or consult a dictionary (100% and 56% for Experienced and Inexperienced readers respectively). The next most frequent behavior for Experienced ESL readers when facing a new English word was to ignore the word (50%). For Inexperienced readers, the second most frequently used strategies were either to ignore the word or to look for contextual clues (31.3% for both). There were no reports of using either linguistic or contextual clues to derive meaning by Experienced ESL readers. Inexperienced readers appeared to try harder to tackle the unknown word either through linguistic cues (25%) or through contextual cues (31.3%).

Table 1

Average frequencies of responses by Experienced and Inexperienced readers in the Metacognitive Interview by category

Questions		Exp.	Inexp.
1.	What is reading?		
	a. Vague/irrelevant	16.7	0.0
	b. Classroom related	41.7	8.8
	c. Word recognition	8.3	0.0
	d. Physiological activity	8.3	25.0
	e. Cognitive act	41.7	75.0
2.	What do you think reading is for?		
	a. Vague or no answer	0.0	6.3
	b. Intrinsic	25.0	12.5
	c. Linguistic	16.7	12.5
	d. Enjoyment/pleasure/fun	50.0	12.5
	e. Functional	91.7	81.3
3.	What do you think makes a good reader?		
	a. Vague or irrelevant	0.0	0.0
	b. Linguistic	0.0	6.3
	c. Procedural	41.7	56.3
	d. Practice/experience	8.3	6.3
	e. Meaning/memory	100.0	93.8
4.	How good a reader do you think you are?		
	a. Very poor	0.0	6.3
	b. Poor	25.0	31.3
	c. Fair	16.7	6.3
	d. Good	50.0	56.3
	e. Very good	0.0	0.0
5.	What would you like to do or could you do that would make you a better reader?		
	a. Vague or irrelevant	8.3	0.0
	b. Linguistic	0.0	0.0
	c. Procedural	16.7	12.5
	d. Practice	66.7	75.0
	e. Cognitive act	16.7	25.0

6A.	When you are reading in ***Chinese*** and come to a word that you don't know, what do you do?		
	a. Ask someone/check a dictionary	58.3	50.0
	b. Ignore it	25.0	43.8
	c. Procedural	0.0	6.3
	d. Decoding	8.3	12.5
	e. Use of context	25.0	25.0
6B.	When you are reading in ***English*** and come to a word that you don't know, what do you do?		
	a. Ask someone/check a dictionary	100.0	56.3
	b. Ignore it	50.0	31.3
	c. Procedural	16.7	12.5
	d. Decoding	0.0	25.0
	e. Use of context	0.0	31.3
7A.	When reading in ***Chinese*** and come to a word that you recognize but don't know the meaning of, what do you do?		
	a. Ask someone/check a dictionary	58.3	68.8
	b. Ignore it	33.3	25.0
	c. Procedural	0.0	6.3
	d. Decoding	0.0	0.0
	e. Use of context	16.7	18.8
7B.	When reading in ***English*** and come to a word that you recognize but don't know the meaning of, what do you do?		
	a. Ask someone/check a dictionary	100.0	81.3
	b. Ignore it	0.0	12.5
	c. Procedural	8.3	6.3
	d. Decoding	0.0	0.0
	e. Use of context	0.0	31.3
8A.	When reading in ***Chinese***, do you ever read something over again? If so, why. If not, why not?		
	a. Never or vague/irrelevant	16.7	6.3
	b. Yes. (But with no reason given)	0.0	6.3
	c. Yes. To learn more	8.3	6.3
	d. Yes. To enjoy more	16.7	6.3
	e. Yes. To understand better.	58.3	75.0
8B.	When reading in ***English***, do you ever read something over again? If so, why. If not, why not?		
	a. Never or vague/irrelevant	8.3	12.5
	b. Yes. (But with no reason given)	8.3	0.0
	c. Yes. To learn more	8.3	6.3
	d. Yes. To enjoy more	0.0	0.0
	e. Yes. To understand better.	66.7	75.0

9A.	What do you do to help you remember what you read in ***Chinese***?		
	a. Vague or irrelevant	41.7	37.5
	b. Procedural	0.0	6.3
	c. Linguistic/textual	0.0	12.5
	d. Practice	50.0	37.5
	e. Cognition	16.7	6.3
9B.	What do you do to help you remember what you read in ***English***?		
	a. Vague or irrelevant	16.7	37.5
	b. Procedural	25.0	18.8
	c. Linguistic/textual	8.3	0.0
	d. Practice	41.7	43.8
	e. Cognition	16.7	0.0
10A.	When reading in ***Chinese***, what do you do to help yourself get the main idea of the material you read?		
	a. Vague or irrelevant	25.0	18.8
	b. Procedural	8.3	12.5
	c. Linguistic/textual	8.3	0.0
	d. Practice	33.3	31.3
	e. Cognition	33.3	43.8
10B.	When reading in ***English***, what do you do to help yourself get the main idea of the material you read?		
	a. Vague or irrelevant	25.0	18.8
	b. Procedural	0.0	12.5
	c. Linguistic/textual	16.7	12.5
	d. Practice	41.7	31.3
	e. Cognition	33.3	31.3
11.	When you are reading for pleasure, what language do you usually read in? In that language, do you read differently when you are reading for pleasure than when you are reading to study?		
	a. No difference	16.7	6.3
	b. Yes. Word level/use of dictionary	16.7	18.8
	c. Yes. Number of times read	25.0	25.0
	d. Yes. Note-taking	25.0	12.5
	e. Yes. Macro-level: speed, inference	58.3	50.0

Similarly, when encountering a familiar word of which the meaning is unknown in the context (Question 7), most readers either sought help from others or from a dictionary whether they were reading in Chinese (58.3% and 68.8% for Experienced and Inexperienced readers respectively) or in English (100% and 81.3% for Experienced and Inexperienced readers respectively). In addition, none of the readers in this study would decode a familiar Chinese or English word of which the meaning is unknown. Inexperienced ESL readers again reported more willingness to look for contextual cues to tackle unfamiliar English or Chinese words than the Experienced group (18.8% vs. 16.7% in reading Chinese and 31.3% vs. 0.0% in reading English).

In Question 8, the majority of readers in this study reported reading Chinese and English texts over and over for better understanding and none of them would re-read English materials for enjoyment. English, being these readers' second language, did not appear to be a language for pleasure reading.

When asked in Question 9 "What do you do to help you remember what you read in Chinese/English?", most readers in both groups considered practice as very helpful in remembering the content, both in reading Chinese and English materials. In reading Chinese, these readers' native language, quite a few readers (41.7% and 37.5% of Experienced and Inexperienced ESL readers respectively) were not aware of any strategies used to help them remember the content. However, in reading in English, more Experienced than Inexperienced ESL readers were aware of their use of procedural practice (25.0% vs. 18.8%), linguistic or contextual clues (8.3% vs. 0.0%), and cognitive acts (16.7% vs. 0.0%) to help them remember what they read. More Inexperienced than Experienced ESL readers gave vague responses (37.5% vs. 16.7%) to this question.

When asked what they did to help get the main idea of the material read (Question 10), most readers reported practices like re-reading and reciting as well as cognitive acts like thinking and forming an internal outline as helpful. Both groups of readers used more linguistic/textual strategies in reading English than in reading Chinese. Also, more Experienced than Inexperienced ESL readers gave vague responses to this question (25.0% vs. 18.8% for both reading in Chinese and in English).

Question 11 asked what language these readers usually read in when they read for pleasure. Only 3 out of the 12 Experienced ESL readers read occasionally in English for pleasure. For the rest of the participants, Chinese was the language for leisure reading. This result is congruent with the result from Question 8 that none of the participants re-read English materials for enjoyment. When reading for pleasure, most readers reported reading faster (58.3% and 50.0% for Experienced and Inexperienced readers respectively) and making fewer regressions (25% for both groups) than when reading for academic purposes.

B. Post-reading Interviews

When asked during the retrospective interview whether they found any parts of the English or Chinese texts difficult, almost all readers reported encountering too many unfamiliar words.

While reading the Chinese manipulated passage, many Experienced readers considered the pseudo words in this passage as either misprints or expressions used in Hong Kong or mainland China. However, three Experienced ESL readers questioned the veracity of the manipulated Chinese article because of the "strange" words used. One Experienced ESL reader commented on her lack of interest in politics and that therefore was a barrier to understanding the article on economy and the China-France relationship. Another reader reported that she could relate strongly to this article because she was interested in politics and the English article on birds did not appeal to her because she did not know much about birds. These two readers agreed that it was the topic rather than the language that caused reading difficulties. One experienced ESL reader reacted emotionally to the situation stated in the article.

Student A:

> I think it is very ironical to have an ambassador from Laos to work in Taiwan. ...This economy-oriented situation in Taiwan is very... I feel very helpless about this situation.

While reading the manipulated English passage, all Experienced ESL readers commented on reading too many difficult words in one article. They however did not seem to falter at these words. One reader showed his tolerance for uncertainty in the text by giving a broader definition of the problematic parts in the text.

Student B:

> (while trying to find out what "urmlews" meant)
> Here, it says "ones that had been unrestrained" obviously they are raised by their parents. So this (referring to "urmlews") is the raising machine in the laboratory.

This reader also ignored many unfamiliar words and identified the structures that bothered him.

Student B:

> It says "like people," but later on there is no mentioning of people. If its purpose is to compare, then how come we did not see any comparison. So

> it's confusing and I think "like people" is not necessary in this sentence.....'Here I first read "a young bird tidly returns once....." then I read it again and realized that it was "once it leaves the nest, a young bird tidly returns."

In reading manipulated English text, only one Experienced reader questioned or challenged the text.

Student C:

> Isn't eating a natural ability? Why do they need to learn how to eat?

For some Inexperienced ESL readers, there were many "eccentric" words in the Chinese article but nobody reported being irritated by these words. Many of them felt very confident that some of the pseudo Chinese words were either journalistic jargon, simplified characters, or misprints. They also reported that these words were of little importance and could be bypassed without affecting comprehension.

In reading the manipulated English passage, the picture became a little different for Inexperienced readers. Nine out of the 16 Inexperienced ESL readers reported encountering many difficult words in reading the English passage, one regarded the content as difficult, and only one reader explicitly reported having difficulty with pronouncing some English words.

According to these readers, the results of these difficulties were nervousness, low self-esteem, comprehension blackout, frustration and pain. For example, being very confident about the "misprinted" pseudo Chinese characters, one reader in this group appeared to be frustrated at the quantity of unfamiliar words in the English passage.

Student D:

> If I had more time I would like to look these words up in the dictionary. Right now the more I read, the more difficult words I found. I just could not go on reading I became very nervous.

Another Inexperienced ESL reader who reported that unfamiliar Chinese words didn't affect comprehension, became very irritated while reading the English passage.

Student E:

> "When it's time to strike out on their own..." I became nervous when I read this...I could not organize ideas here and there were so many new words. I felt like an idiot who kept reading without understanding. Because I got

> lost in the beginning, I lost my confidence and sense of achievement I would definitely look them up if I were reading at home.

Other readers, however, attributed the difficulties to the lack of content schema besides vocabulary.

Student F:

> This string of words (eagles, hawks, refirk...) is a barrier in reading. I read them aloud but still didn't know what they meant. I major in social science and don't know much about natural science, therefore this article is difficult to me.

Five readers in this group reported that they would check all the difficult words in the dictionary if they were reading at home.

The strategies reported by participants were tabulated and are presented in Table 2. Table 2 shows that for Experienced ESL readers guessing was the strategy most commonly used in reading English and Chinese texts. In reading Chinese, Experienced ESL readers relied more heavily on contextual cues (27%), though word analysis was the next most frequently used strategy (24%).

Table 2

Percentages for use of strategies by the two groups of readers when reading manipulated Chinese and English passages

Group Passages	Strategies (%)							
	1	2	3	4.1	4.2	4.3	4.4	5
Exp Chinese	18	9	6	9	0	24	27	6
Exp English	13	3	13	7	10	30	23	0
Inexp Chinese	56	6	0	18	0	12	9	0
Inexp English	16	2	3	0	14	27	33	5

However, in reading English, the dependence was more on word analysis (30%) and less on context (23%). The Inexperienced ESL readers seemed to be more tolerant of ambiguity in reading Chinese: 56% of the time they ignored the problematic parts and carried on with their reading. In reading the English passage, they also seemed to rely more on context (33%) than on word analysis (27%).

In summary, participants' retrospective reports showed that most readers were less bothered by difficult words in their native language than in an L2. When confronted with unfamiliar words, frustration, nervousness and low self-esteem appeared to be common feelings for Inexperienced ESL readers. The Inexperienced ESL readers were more tolerant of ambiguity in Chinese with most of the uncertainty in the text ignored. However, in reading English, the same group of readers, the Inexperienced ESL readers, relied on context when faced with difficulties. Experienced ESL readers resorted to context in reading Chinese and relied on lexical features in reading English when encountering difficulties. Besides vocabulary, readers' personal interests in topics also contributed to the ease/difficulty of reading.

The observations made in this study of Chinese readers reading in Chinese and English are consistent with Chen's (1990) study of Chinese readers in America. Both studies found that readers were more confident reading in their native language and rely more on local strategies when reading in English.

Discussion

The different attitudes towards the purposes of reading reflected by Experienced and Inexperienced readers is worth further exploration. Mature readers, fourth year university students, considered reading as both functional and enjoyable, whereas first year university students regarded reading as basically function-oriented, i.e., to gain knowledge and to know the writer's thoughts. The exam-oriented high school instruction and the pressure of the university entrance exam, which most Inexperienced readers had just experienced, may have shaped these readers' concept of reading in general.

Participants' responses to Questions 6 and 7 showed that, unlike the Inexperienced readers who looked for linguistic and contextual clues, Experienced readers didn't seem to expend much effort on difficult words. However, when it came to remembering the content or getting main ideas (Questions 9 and 10) in reading English materials, Experienced readers appeared to venture more guesses and were more aware of their use of strategies. Obviously, meaning-getting played a more important role than sorting difficult words out in Experienced readers' reading process.

Bruinsma (1990) interviewed elementary school students to tap their cognitive awareness in reading, and found that "practice makes perfect" was an axiom of the majority of these students. In other words, these young students believed that good readers practiced reading more and were experienced in reading, they also believed that to become a good reader, one had to practice more (i.e. read more). The Chinese university students who participated in this study also considered "practice" as an important avenue to becoming a better reader. However, when asked what made a good reader, these students chose meaning-getting and speed/accuracy as the main attributes of a good reader. Practice/experience didn't receive as much weight in this case as in Bruinsma's study.

Decoding (phonic or structural analysis), a common strategy for young participants in Bruinsma's (1990) study, did not appeal to participants in this study either in reading Chinese or English. However, Inexperienced ESL readers did suggest the use of decoding unfamiliar words relatively more frequently than Experienced ESL readers both in reading Chinese and in reading English. This is in line with the results of Winser's (1988) study which found less proficient native English readers to be more reliant on sounding out as a strategy to get at meaning. The adult readers in this study tended to turn to other people or resort to a dictionary when encountering difficult words in reading, be it reading in an L1 or an L2. This result reconfirmed a commonly observed phenomenon that Chinese readers were dictionary-dependent and accuracy-oriented in reading.

Compared with responses to questions 1 to 3, questions 9 and 10, which were on strategies to remember the content or to get the main idea, received more vague or irrelevant responses. This showed readers' lack of metacognitive awareness of their reading behaviors in getting meaning from or remembering the text. This became more obvious when readers reflected on how they read in their native language, Chinese. Probably reading in Chinese has become an internalized process and the end result, comprehension, has been taken for granted therefore no analytical thinking surfaced when asked. Also one traditional concept in Chinese education is that "memorization is the key to comprehension." If this concept has been rooted in these readers' minds in their previous schooling in Chinese, then, to these readers, no strategies are necessary to get at the content of the Chinese materials read.

That Experienced ESL readers showed more awareness of their metacognitive skills is in line with the claim that though all language learners use strategies, the more effective students use them more consciously and more frequently than the less effective learners (Oxford, Crookall, Cohen, Lavine, Nyikos and Sutter 1990). The claim that the ability to effectively use metacognitive skills is a crucial component of skilled reading (Geva and Ryan 1985, Singer and Donlan 1982) has found support in this study.

Pedagogical implications

Though interview results only reflected interviewees' own perspectives of what they did and it is likely that they have guessed what the investigator wanted and responded accordingly, the results of this study have shed light on differences in

L1 and L2 learning. Some teaching implications have emerged from this study. Awareness of using strategies to get at meanings in reading an L2 appeared to distinguish Experienced ESL readers from Inexperienced ESL readers in this study. The finding that inexperienced readers appeared to be more aware of strategies in tackling words than in dealing with global comprehension reveals the need to emphasize the skills for holistic reading.

Increased metacognitive awareness has been proved to lead to better use of reading strategies by children (Paris, Cross and Lipson 1984) as well as by L1 and L2 learners (Carrell 1989). And training has been demonstrated to result in an increase in monitoring reading comprehension in children as well as adults (Grabe and Mann 1984). Therefore, it is advisable to enhance students' awareness of their own reading processes and develop their ability in selecting appropriate strategies through instruction. Also, future instructional emphasis should be placed on providing a more supportive and non-threatening environment to help ESL students develop a better understanding of the reading process.

References

Chen, D. (1990). Effects of formal schema and metacognition on reading comprehension of Chinese and English expository texts by Chinese readers. Unpublished doctoral dissertation, University of Michigan.

Barnett, M.A. (1988). Teaching reading strategies: How methodology affects language course articulation. *Foreign Language Annals*, 21:109-119.

Bruinsma, R. (1990). Learning to ride a bike and learning to read: Children's conception of reading. *Australian Journal of Reading*:13:114-127.

Carrell, P.L. (1989). Metacognitive awareness and second Language reading. *The Modern Language Journal*, 73:121-134.

Devine, J. (1984). ESL reader's internalized models of the reading process. In J. Handscombe, R. Orem, & B. Taylor (eds.), pp. 95-108 *On TESL '83*, Washington, DC: TESOL.

Geva, E. & Ryan, E.B. (1985). Use of conjunctions in expository texts by skilled and less skilled readers. *Journal of Reading Behavior*, 17:331-346.

Goodman, Y.M., Watson, D.J. & Burke, C.L. (1987). *Reading miscue inventory alternative procedures*. New York: Richard C. Owen Publishers.

Grabe, M. & Mann, S. (1984). A technique for the assessment and training of comprehension monitoring skills. *Journal of Reading Behavior*, 16:131-144.

Kern, R.G. (1988). *The role of comprehension strategies in foreign language reading*. Unpublished dissertation, University of California, Berkeley.

McNeill, J.D. (1987). Metacognition in reading comprehension. In *Reading comprehension: New directions for classroom practice* (2nd edit.) pp. 91-105. Glenview, IL: Scott, Foresman and Company.

Oxford, R., Crookall, D., Cohen, A., Lavine, R., Nyikos, M. & Sutter, W. (1990). Strategy training for language learners: Six situational case studies and a training model. *Foreign Language Annals*, 22:197-216.

Padrón, Y., Knight, S.L., & Waxman, H.C. (1986). Analyzing bilingual and monolingual students' perceptions of their reading strategies. *The Reading Teacher*, 39:430-433.

Paris, S.G., Cross, D.R. & Lipson, M.Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology*, 76:1239-1252.

Singer, H. & Donlan, D. (1982). Active comprehension: Problem-solving schema with question generation for comprehension of complex short stories. *Reading Research Quarterly*, 17:166-186.

Winser, B. (1988). Readers getting control of reading. *Australian Journal of Reading*, 11:257-268.

Appendix A:

Metacognitive Interview Questions

1. What is reading?

2. What do you think reading is for?

3. What do you think makes a person a good reader?

4. How good a reader do you think you are?

   Very poor__ Poor___ Fair___ Good___ Very good____

5. What would you like to do or could you do that would make you a better reader?

6. (a) When you are reading in Chinese and you come to a word that you don't know, what do you do?

(b) When you are reading in English and come to a word you don't know, what do you do?

7. (a) When you are reading in Chinese and come to a word you recognize but don't know the meaning of, what do you do?

   (b) When you are reading in English and come to a word you recognize but don't know the meaning of, what do you do?

8. (a) Do you ever read something over again in Chinese? Why or why not?

   (b) Do you ever read something over again in English? Why or why not?

9. (a) What do you do to help you remember what you read in Chinese?

   (b) What do you do to help you remember what you read in English?

10. (a) What do you do to help yourself get the main idea of the reading material when you are reading in Chinese?

    (b) What do you do to help yourself get the main idea of the reading material when you are reading in English?

11. When you are reading for pleasure, do you usually read in Chinese or in English?

12. When you are reading for pleasure, do you read differently than when you are reading to study? (If yes) What do you do differently when reading for pleasure?

Appendix B:

Metacognitive Interview Coding Taxonomy

1. What is reading?

   Categories:

   a. Vague or irrelevant
   e.g. Reading is "just read," and "when you read." "I don't know."

   b. Classroom/object related procedure
   e.g. "Reading a book/a word/a story." "Reading is reading something made of words."

   c. Word recognition (decoding)

e.g. "Sounding out letters/words." "Looking at words and saying them." "Reading is learning new words."

d. Physiological activity
e.g. "Reading is an exercise of eyes/brain."

e. Cognitive act (meaning)
e.g. "Reading is when you understand and learn knowledge." "Reading is to remember the content."

2. What do you think reading is for?

Categories:

a. Vague or no answer
e.g. "I don't know."

b. Intrinsic
e.g. "To become good at reading." "To become a better person." "To stimulate thinking." "To become smarter."

c. Linguistic
e.g. "To learn more words." "To learn better expressions." "To improve language skills."

d. Enjoyment/pleasure/fun
e.g. "Reading is for fun/for recreation." "I like it." "I read when I am bored."

e. Functional
e.g. "To know things." "To gather data." "To learn new information." "To understand the writer's thoughts."

3. What do you think makes a good reader?

Categories:

a. Vague or irrelevant

b. Linguistic
"A good reader knows a lot of words/expressions." "A good reader usually writes better."

c. Procedural
e.g. "A good reader reads fast/makes no mistakes."

d. Practice/experience
   e.g. "A good reader is one who reads every day (a lot)/one who is old."

e. Meaning/memory
   e.g. "A good reader really knows what's going on when talking about the story." "A good reader has good comprehension/reads critically/can catch main idea."

4. How good a reader do you think you are?

   Categories:

   a. Very poor
   b. Poor
   c. Fair
   d. Good
   e. Very good

5. What would you like to do or could you do that would make you a better reader?

   Categories:

   a. Vague or irrelevant

   b. Linguistic
      e.g. "Increase vocabulary."

   c. Procedural
      e.g. "Read slower/faster." "Read with good expression."

   d. Practice
      e.g. "Read more books." "Learn from others." "More training."

   e. Cognitive act
      e.g. "Think more in my head as I read." "Concentrate more." "Meditate before I read.

6. When you are reading in Chinese/English and come to a word that you don't know, what do you do?

   Categories:

   a. "Ask someone." / "Check a dictionary."

b. Ignore it
e.g. "Skip it." "Read on."

c. Procedural
e.g. "Underline/mark/write down the word."

d. Decoding (phonic/structural analysis)
e.g. "Try to pronounce the word." "Find out its part of speech/function in the sentence."

e. Use of context (cognitive strategies)
e.g. "Go to the previous/following sentences and look for clues." "Skip it for now and wait for further clues." "Guess the meaning."

7. When you are reading in Chinese/English and come to a word that you recognize but don't know the meaning of, what do you do?

Categories:

a. "Ask someone." / "Check a dictionary."

b. Ignore it
e.g. "Skip it." "Read on."

c. Procedural
"Underline/mark/write down the word." "Read more times."

d. Decoding (phonics and/or structural analysis)
e.g. "Try to pronounce the word." "Find out its part of speech/function in the sentence."

e. Use of context (cognitive strategies)
e.g. "Go to the previous/following sentences and look for clues." "Skip it for now and wait for further clues." "Guess the meaning."

8. When reading in Chinese/English, do you ever read something over again? Why or why not?

Categories:

a. Never or vague or irrelevant

b. Yes. (But with no reason given.)

c. Yes. To learn words/expressions/structures.

d. Yes. To enjoy more.

e. Yes. To understand better.

9. What do you do to help you remember what you read in Chinese/English?

Categories:

a. Vague or irrelevant
e.g. "Nothing special."

b. Procedural
e.g. "I translate it."

c. Linguistic/textual
e.g. Apply a particular skill like phonics, slow reading (word by word or sentence by sentence). Read with contextual aids (pictures, titles etc), or read key words.

d. Practice
e.g. "Re-reading/reciting/trying to memorize it." "Make a note/write an outline/paraphrase it in the margin.

e. Cognition
e.g. "Think about what I had just read." "Form an outline in my mind." "Focus on topic sentence." "Think of related material."

10. When reading in Chinese/English, what do you do to help yourself get the main idea of the material you read?

Categories:

a. Vague or irrelevant
e.g. "Nothing special."

b. Procedural
"Translate it." "Read others' critiques or comments."

c. Linguistic/textual
e.g. Apply a particular skill like phonics, slow reading, read with contextual aids.

d. Practice
e.g. "Re-reading/reciting/trying to memorize it." "Make a note/write an outline/paraphrase it in the margin."

e. Cognition
e.g. "Think about what ! had just read." "Form an outline in my mind." "Focus on topic sentence." "Think of related material."

11. When you are reading for pleasure, what language do you usually read in? In that language, do you read differently when you are reading for pleasure than when you are reading to study?

Categories:

a. No difference

b. Yes. Word level/use of dictionary

c. Yes. Different number of times read

d. Yes. Note-taking

e. Yes. Macro-level: speed, inference

Appendix C:

Post-Reading Interview Coding Form

1. Read on/skip/ignore
Responses like "I just keep reading," or "This is not important so I just skip it," were coded in this category.

2. Inference
Responses like "I try to think of related subject matter" were coded in this category.

3. Re-read
This was coded when readers reported "I read it again."

4. Guess (assign a temporary meaning)
This is a general category (4.1) used to code reports like "I just guess." Where readers used the word "guess" but with specific guessing strategies, these were coded in the following way: (4.2) guess from neighboring words, (4.3) guess from prefixes, suffixes and parts of speech, and (4.4) guess from context.

5. Mark the text
Responses like "I mark the words I don't know" and "I underline the whole parts that are unclear to me" belonged to this category.

THE IMPACT OF ILLUSTRATIONS AND CULTURAL SCHEMATA ON HONG KONG PUPILS' READING COMPREHENSION AND RECALL OF TEXT

Terry Dolan

Introduction

A fundamental assumption underlying the teaching of reading in schools is that pupils will be able to use their reading ability as a vehicle for learning (Lunzer and Gardner 1979). A second is that, where appropriate, they will draw upon previous reading to solve problems associated with current studies. Hence, textbooks are not only regarded as repositories of essential information and subject matter, but as sources of exercises required for mastering subject disciplines. It is assumed that some of the knowledge in school texts is so important that it needs to be learnt verbatim, whilst pupils need only extract the essential messages or themes from other pages. Over sixty years ago, Bartlett (1932) showed that readers seldom commit to memory each word they read in a passage verbatim. Instead, on the basis of what they have understood already, they mentally précis units of text and keep these summaries in mind as they seek the overall structure or gist. Comprehension is hence governed by what readers already know about the topic, what they have extracted from the passage so far and how effectively they link elements of meaning appearing as the passage unfolds.

Bartlett also showed how comprehension is influenced by cultural background, with readers from one ethnic group failing to grasp the significance of passage detail based on unfamiliar cultures, but effortlessly assimilating familiar schemata from their own cultural experiences. Tested subsequently to see how much they could remember, all subjects recalled the gist of passages more readily than the details and all groups best recalled information relating to their own cultural background.

The significance of such findings will not escape the teacher of English as a second language: pupils reading texts with content based on the culture of speakers of the target language may recall and understand information less well than if it was framed against culturally familiar schemata. At present, it is usual for reading practice to be given using texts based mainly on schemata belonging to indigenous speakers of the target language, for this both broadens horizons and facilitates insight into their world. However, in view of Bartlett's findings, it is worth considering giving pupils reading practice using second language texts featuring more familiar subject matter.

As they read, good readers will temporarily disregard text signals whose importance is not immediately apparent and read on in search of the overall meaning (Meyer 1984). In contrast, poorer readers, especially of a passage with unfamiliar content, are inclined to latch on to any known item of vocabulary in the hope that it will give some idea about passage content and structure (Ohlhausen and

Roller 1988, Kintsch 1987). Will readers of a difficult text in a foreign language do the same, clutching at any clue in the hope that it will unlock the message the author has written? If this is indeed the case, then pupils' comprehension may be swayed by schemata which bear on their own rather than on a foreign culture, whether or not such items are pertinent to the essential theme of the passage.

Illustrations accompanying text are usually intended to clarify, complement or add to the written information. However, unlike the situation with text where the amount of information conveyed can be controlled fairly precisely by the author, the amount of information in an illustration can be vast and its interpretation hard to regulate (Levie 1987). Thus, while readers may draw inferences from illustrations which support or supplement the meaning in a text, equally importantly they may also form mental representations at odds with or even contradicting the author's intended meaning. Levin et al. (1987) suggest that all readers will in some degree resort to consulting accompanying pictorial information when faced with conceptually demanding or incoherently written text. Hence one might expect poor readers or those who lack confidence in their understanding of the language deliberately to seek guidance from the illustrations in a text.

Kintsch (1974) and van Dijk and Kintsch (1983) propose that the information presented in a text may be analyzed in terms of 'propositions', 'atoms of meaning' which together constitute the overall meaning or 'macrostructure'. Reading comprehension, it is suggested, involves representing to oneself the meanings one thinks reside in text. Good readers are skilled at spotting good prospects to hold in mind and carry forward to forthcoming sentences. If comprehension breaks down they usually resort to one of two options: either they press ahead hoping that things will be clarified by subsequent information, or they back-track seeking to identify the point where understanding faltered (Kintsch 1987). Sanford and Garrod (1981), contesting Kintsch's assertion that the macrostructure of text is assembled from how it is propositionally analyzed by the reader, propose that texts are interpreted through a complex series of representations resulting in 'messages' rather than propositional lattices.

Regardless of this dispute, none of the above theorists would challenge the view held by many teachers that pupils listening to or reading a foreign language have a constant fear that the information being carried forward during comprehension is unsound. Language teachers know that even slight suggestions that information is not properly understood may cause pupils' confidence to waiver.

Research into the contribution of illustrations to reading comprehension has highlighted how they sometimes compete with words as sources of information. Samuels (1970) showed how illustrations often seduce the poor reader into relying on the illustration rather than the words to access meaning in sentences. Schallert (1980), researching the function of illustrations in extended text, concluded that they undoubtedly have a positive effect on the comprehension both of narrative and expository prose. After an exhaustive review, Levie and Lentz (1982) concluded that most research has shown that reading comprehension is facilitated by illustrations amplifying or supplementing information in the text, but that cosmetic or incidental illustrations have little impact. In an intriguing study, Read and Barnsley (1977) presented adult Canadians with versions of a reader they had last seen in elementary school. One version contained the text alone; one showed the illustrations and nothing else; and one contained both. Virtually nothing was cued either by the text alone or pictures alone, but when the two were presented together a surprising amount of the story was instantly recalled.

In a key study involving pupils reading in a foreign language, Jahoda et al. (1976) examined the effects of illustrations in culture-free expository text with samples of first and second language pupils of secondary school age. They showed convincingly that the advantage with all ethnic groups was significantly in favour of pupils who had read text with illustrations. The pupils sampled by Jahoda et al. were fairly proficient in English, and one wonders what would have happened if for some groups the illustrations had been more of a life-line for comprehension than the text itself, as is often the case with readers with poor English.

Denburg (1977) proposes that there is a trade-off between written and pictorial information, and that efficient readers of a straight-forward text will note but not scrutinise illustrations. If they meet an illustration which seems incongruous to the text they will register this but press ahead. On the other hand, poor readers detecting an inconsistency between text and illustration instinctively assume that it is their own poor grasp of English which is the source of the confusion. Such uncertainty and self-doubt is familiar to foreign language teachers and many will repeatedly give pupils exercises involving rote learning to ensure that vital elements of the target language can be called upon at will to facilitate communication and further learning.

With the above points in mind the following experiment was carried out, subjects being drawn from the United Kingdom and from a Hong Kong school in which the curriculum is delivered principally in English. Pupils were given different versions of a text in English and asked to read it silently in preparation for a memory test to be given some time in the future. It was anticipated that the children's ability to assimilate and subsequently recall what they had read would reflect their facility with English, their response to the text and whether they had consciously sought guidance from the illustrations.

Method

Textual Material

An adventure story was written in booklet form around four black and white illustrations, each giving support but not adding materially to the story line. In a second booklet, illustrations showing the main characters in the story in settings incongruous to the story replaced the original pictures. A third booklet contained the text alone. Two versions of each booklet were then prepared: one with western characters, background and scenes, the other with a Hong Kong background and characters with Chinese names. Thus a total of six booklets were prepared. The text was validated with Form Three secondary pupils in Hong Kong, who read it drawing attention to any wording hard to understand. Amendments were made until the text was judged easy to follow.

Subjects

The Hong Kong sample was drawn from Form Three classes in an Anglo-Chinese grammar school admitting children predominantly from Bands One and Two. The United Kingdom sample was drawn from Year Three classes in an urban comprehensive school in Nottingham whose population includes the full social and intellectual range. Children were randomly allocated to groups and no child needed to be excluded on the grounds of poor reading ability. Table 1 summarises the distribution of subjects to the various experimental groups.

Table 1

Distribution of Subjects to Groups

		CIT	ICT	TA	Total
HK group	UK version	24	23	25	72
HK group	HK version	42	42	37	121
UK group	UK version	13	13	12	38
UK group	HK version	10	9	11	30
Total		89	97	85	261

*Note: CIT = Congruously Illustrated Text
ICT = Incongruously Illustrated Text
TA = Text Alone

Procedure

The writer was assisted by two experienced teachers, one from Nottingham and one from Hong Kong. Both were accomplished researchers, were well known to the pupils and each held a recent higher degree. Pupils completed the initial reading in normal classroom conditions, having been told in their mother tongue that they had been chosen to take part in an important experiment to discover how much people remember of what they read. They were handed a colour-coded test booklet, informed they would be tested in three weeks' time and given as long as they wished to read and digest the story. Taking the task seriously, pupils from England took about twenty minutes to complete the task and pupils from Hong Kong about ten minutes longer.

Three weeks later, the pupils were tested as whole classes during regular English lessons. Given paper, they were asked to write down all they could recollect of the story but not to add their names. The British pupils took about twenty-five minutes and the Hong Kong pupils between two and thirty-five minutes to write down the story. After the papers had been handed in, the full purpose of the experiment was disclosed to the pupils who were then asked about the effects of the schematic background of the story and the illustrations. Even though some children could recall little, most were willing to talk about the task and seemed genuinely interested.

Scoring

Scripts were marked for recall of detail and recall of gist. Scoring for detail was based on procedures by Meyer (1975) and Lunzer and Byron (1981), and consisted of segmenting the text into text units, each of which attracted one mark. Scoring for gist took account of procedures developed by Riding and McQuaid (1977) and consisted of seeing how closely each script included reference to key developments in the story, whether or not identical wording had been used. Placing events and developments in their correct story sequence also earned marks. The scoring procedure earned the approval of a panel of experienced public examination markers and psychologists. Each script was marked blind and independently by members of a team of judges and any discrepancies sorted out. Scripts from subjects who had received ICT booklets were examined for information which had appeared in the illustrations but not in the text.

Results

The main interest in the present paper is in the results gained by the Hong Kong subjects, but the performance of the United Kingdom pupils is summarised below for reference.

As expected, the British pupils were significantly better than their Hong Kong counterparts at recalling information. With the British group alone, multivariate analysis of variance revealed no significant main effects or interactions, either for recall of detail or of gist. Post-test discussions suggested that performance had been principally influenced by the reading of the text rather than by the illustrations. Nevertheless, there was a slight but non-significant tendency to do best with texts with 'British' schemata, and with CIT material. Many children who had had the ICT stories said they had noticed that the pictures did not match the story but had ignored this. Pupils reading the 'Hong Kong' schematic material had been generally attracted by the unusual names and location of the story, and many were able to recall the Chinese names perfectly. An interesting point is that the British children had not had such a test before and welcomed the opportunity to test their memory for textual material.

Turning to the performance of the Hong Kong group, the first point to note is that many children failed to recall either the detail or gist of the story. Thus, attempts to apply multivariate analysis of variance floundered due to a lack of homogeneity in the data. In order to effect a cogent analysis, subset analyses were carried out using one-way analysis of variance and a formula accepting cells of unequal size (Youngman 1976).

The influence of the cultural (schematic) background of the passage on pupils' recall of detail and gist is summarised in Table 2. It is very clear that the children were much better at recalling both the gist and the details of the story based on their own cultural background than the story involving United Kingdom schemata and background information.

Table 2

Effects of Cultural Schemata on Recall

	Detail		**Gist**	
	mean	**s.d.**	**mean**	**s.d.**
UK booklets	3.12	3.00	2.74	2.19
HK booklets	7.78	8.86	4.31	2.51
between-group testing	t=4.29 (p<.01)		t=4.40 (p<.01)	

Table 3 summarises the effects of type of text on performance. As may be seen, pupils reading the CIT booklets did almost twice as well as their classmates in the other two groups in terms both of recalling detail and penetrating the gist of the story. In fact, the TA group recalled slightly more detail than the ICT group, but less well in terms of remembering the gist.

Table 3

Effects of Text Type on Recall

	detail		**gist**	
	mean	**s.d.**	**mean**	**s.d.**
CIT (n = 66)	8.80	9.69	4.56	2.51
ICT (n = 65)	4.55	6.67	3.49	2.37
TA (n = 23)	4.66	4.53	3.08	2.43
One-way ANOVA	F-ratio 7.0119 (p<.01)		F-ratio 6.2399 (p<.01)	

Table 4 summarises performance when the effect on recall of type of text is analyzed alongside the schematic background of the story. The results could not be more clear: pupils reading booklets based on their home background did better in every instance than their peers reading booklets based on the United Kingdom. Generally speaking, pupils did best on recalling CIT material; they recalled the TA booklets better than the ICT material for the United Kingdom based stories, in terms both of content and detail; and did slightly better on texts with Hong Kong based illustrations even when they did not fit the story.

Table 4

Effects of Text Type and Schematic Background

	Detail		**Gist**		
	mean	**s.d.**	**mean**	**s.d.**	**N**
CIT - UK schemata	3.50	3.10	3.54	2.12	24
CIT - HK schemata	11.83	10.81	5.14	2.53	42
ICT - UK schemata	2.70	2.73	2.22	2.08	23
ICT - HK schemata	5.75	7.97	4.19	2.22	42
TA - UK schemata	3.16	3.11	2.44	2.14	25
TA - HK schemata	5.68	5.03	3.51	2.52	37
One-way ANOVA	F-ratio 8.4062 (p<.01)		F-ratio 6.8635 (p<.01)		

In order to examine more precisely the relationship between gist and detail recall and to rule out uncontrolled between-group differences, it was decided to carry out a regression analysis in which the gist scores of the entire Hong Kong sample, irrespective of sub-group, were used to predict the detail score for each subject. The differences between predicted and obtained scores were transformed into residual scores then standardised around a mean of zero. The results are shown in Table 5, where a negative score signifies worse than anticipated performance and a positive sign signifies better.

On texts with United Kingdom schema, the children did worse than their overall pattern of gist scores would suggest. In particular, the children who had read the CIT story remembered detail less well than their memory for gist would suggest. In contrast, with the Hong Kong based material, children who had read the CIT story recalled text detail very well, especially in comparison with pupils in the other groups. It is interesting to note that subjects who had read ICT material displayed relatively poor recall of detail.

Table 5

Residual Analysis: Schemata by Text Type

	Residual scores	
	mean	s.d.
CIT - UK	-2.14	3.93
ICT - UK	-0.08	2.78
TA - UK	-0.10	3.51
CIT - HK	2.72	7.62
ICT - HK	-1.48	4.74
TA - HK	0.09	4.21
One-way ANOVA	F-ratio 3.9386 p<.01	

Discussion

Post-test interviews with the children cast some light on why the two national groups performed so differently. Whereas the British children found the testing of their memory an unusual and interesting exercise, the Hong Kong pupils regarded such testing as routine in their run up to examinations. In fact, they claimed they had had several such tests in the interval between the initial and follow-up experimental sessions. They welcomed the presence of illustrations in the story, whereas the British children considered them unimportant. Aware that their task was to memorize the text, most British children said they had neither studied the pictures nor been put off by pictures which did not appear to fit the story line. Having decided the pictures were unlikely to be significant, they simply dismissed them. In contrast, many of the Hong Kong children admitted to being puzzled by the incongruous pictures, to presuming that the pictures must be relevant and that somehow their English comprehension was at fault.

Although the British children took little heed of the illustrations, most admitted glancing at them. Studies by Goodykoontz (1936), Peeck (1974, 1985) and Dolan E. (1988) also found that secondary pupils do not seem to take any discernable notice of illustrations when reading stories in their mother tongue. Hence, it is interesting to note that the Hong Kong group said they had deliberately consulted the illustrations, considering them vital adjuncts to the text. This comprehension strategy is of course taught systematically in reading kits such as the SRA but is a technique rarely used consciously by skilled and confident readers of undemanding text (Fawcett 1979). Jahoda et al. (1976) report that the only illustrations having any impact on the second language readers in their study were those portraying material not present in the text but, as mentioned earlier, all the children in their experiment were quite proficient in English.

Returning to the present study, both national groups said that during the recall session they were aware of the pictures providing cues and clues about content. Although the presence of 'English' and 'Chinese' pictures and schemata was dismissed as being beside the point by the British children, many pupils said that the Chinese names were so unusual that they had deliberately memorized them. Contrast this with the many Hong Kong pupils who frankly confessed that they usually 'skip' foreign names in stories. In fact, some said their interest waned instantly on seeing the English names and pictures. One has to remember that the text had been developed chiefly with Hong Kong pupils, so the British groups had found it quite simple and might have had recourse to the pictures had the language been more demanding. Hence, the TA British group said they would not have benefited from pictures in the text, whereas the Hong Kong group said they had found it a chore to face unbroken text and would have welcomed pictures.

The incongruous pictures had little significant impact on the British children. However, a number of Hong Kong children actually wove their story around the illustrations, for example writing about "robbers and thieves" rather than "kidnappers". One child wrote, "the police catch the man who had not hair", reference to a picture showing a bald man, a point not actually mentioned in the story. The only evidence of a British child being misled came from a boy who wrote, "the police used a spotter plane", reference to an incongruous illustration showing the plane in the air, whereas the kidnappers in the actual story were apprehended before they had time to take off.

A serious point to ponder is that, whereas the British children welcomed the challenge of sorting out for themselves the meaning residing in the text, the Hong Kong children gave the impression that they would sooner have relied upon the guidance of their English teacher, guidance which they indicated is usually forthcoming. The insinuation is that some Hong Kong teachers seem to doubt the capacity of their pupils to fathom unaided the main points and gist contained in extended English prose. They thus resort to 'duck-feeding', ensuring uniformity of knowledge coverage but doing little to build up pupils' confidence in their ability to work independently.

If it is indeed the case that second language learners who lack confidence in their language proficiency will seek refuge in illustrations when they experience problems in understanding text, one ought to be aware that, although illustrations may facilitate matters, it is unwise to presume that understanding has been arrived at through verbal processing whenever visual aids to comprehension are also present. However, one would not argue that illustrations serve no useful purpose. Far from it! They captivate interest, add appeal to reading material and offer fascinating insights into salient aspects and unfamiliar schemata. But this is not to suggest that it is always essential, or even desirable, at the practice stage for readers to encounter a diet of texts based on the target culture alone. Rather, beginning readers of a second language will benefit from exposure both to texts containing target language schemata and texts featuring schemata and content familiar to their own cultural background.

References

Artlett, C. (1932). *Remembering: a Study in Experimental and Social Psychology.* Cambridge: Cambridge University Press.

Denburg, S.D. (1977). The interaction of picture and print in reading instruction, *Reading Research Quarterly,* 12:176-189.

Dolan, E. (1988). Memory for Text. Unpublished M.Ed. Dissertation, University of Nottingham.

Fawcett, R. (1979). Reading laboratories. In E. Lunzer & K. Gardner (eds.), *The Effective Use of Reading,* London: Heinemann.

Goodykoontz, B. (1936). The relation of pictures to reading comprehension, *Elementary English Review,* 13:125-130.

Jahoda, G. Cheyne, W.M. Deregpwski, J.B. & Collingbourne, R. (1976). Utilization of pictorial information in classroom learning: a cross-cultural study, *AV Communication Review,* 24:295-315.

Kintsch, W. (1974). *The Representation of Meaning in Memory.* Hillsdale: Lawrence Erlbaum Associates.

Kintsch, W. (1987). Contributions from cognitive psychology. In T.J. Tierney, P.L. Anders & J.N. Mitchell (eds.), *Understanding Readers' Understanding.* Hillsdale: Lawrence Erlbaum Associates.

Levie, W.H. (1987). Research on pictures; a guide to the literature. In D.M. Willows & H.A. Houghton (eds.), *The Psychology of Illustration: Volume I, Basic Research.* New York: Springer-Verlag.

Levie, W.H. & Lentz, R. (1982). Effects of text illustrations: a review of research, *Educational Communication and Technology Journal,* 30:195-232.

Levin, J.R. Anglin, G.J. & Carney, R.N. (1987). On empirically validating functions of pictures in prose. In D.M. Willows & H.A. Houghton (eds.) *The Psychology of Illustration: Volume I, Basic Research.* New York: Springer-Verlag.

Lunzer, E. & Byron, N. (1981). Suiting the task to the text. Paper read at the 26th IRA Convention New Orleans: IRA.

Lunzer, E. & Gardner, K. (eds.) (1979). *The Effective Use of Reading.* London: Heinemann.

Meyer, B.J.F. (1975). *The Organization of Prose and its Effects on Memory,* New York: North-Holland.

Meyer, B.J.F. (1984) Text dimension and cognitive processing. In H. Mandl, N.L. Stein & T. Trabasso (eds.) *Learning and Comprehension of Text.* Hillsdale: Lawrence Erlbaum Associates.

Ohlhausen, M.M. & Roller, C.M. (1988). The operation of text structure and content schemata in isolation and in interaction. *Reading Research Quarterly,* 23:70-88.

Peeck, J. (1974). Retention of pictorial and verbal content of a text with illustrations. *Journal of Experimental psychology,* 66(6):880-888.

Peeck, J. (1987). The role of illustrations in processing and remembering illustrated text. In D.M. Willows & H.A. Houghton (eds.) *The Psychology of Illustration: Volume I, Basic Research.* New York: Springer-Verlag.

Read, J.D. & Barnsley, R.H. (1977). Remember Dick and Jane? Memory for elementary readers. *Canadian Journal of Behavioral Science,* 9:361-370.

Riding, J. & McQuaid, D.G. (1977). The effect of detail importance and passage structure on the retention of prose by twelve-year-old children. *Educational Studies,* 3(2):145-151.

Samuels, S.J. (1970). Effects of pictures on learning to read. *Review of Educational research,* 40:397-408.

Sanford, A.J. & Garrod, S.C. (1981). *Understanding Written Language.* Chichester: Wiley.

Schallert, D.L. (1980). The role of illustrations in reading comprehension. In R.J. Spiro, B.C. Bruce & W.F. Brewer (eds.) *Theoretical Issues in Reading Comprehension.* Hillsdale: Lawrence Erlbaum Associates.

Van Dijk, T.A. & Kintsch, W. (1983). *Strategies of Discourse Comprehension.* New York: Academic Press.

Youngman, M.B. (1976). *Programmed Methods for Multivariate Data.* Nottingham: Nottingham University School of Education.

WORD PRINTED FREQUENCY/FAMILIARITY AND STRUCTURE COMPLEXITY EFFECTS ON L1 AND L2 WORD RECOGNITION PROCESSES IN CHINESE

Yilin Sun

Introduction

It is widely recognized that the interpretation of the effect of word frequency is central to many models of word recognition. It is generally considered that morphemic and lexical recognition units vary in accessibility according to their familiarity. Research has indicated that, first, words with regular pronunciations yield faster naming latencies than words with irregular pronunciations (exception words) when the words are low in frequency. Higher frequency regular and exception words yield similar naming latencies (Seidenberg 1985b, Seidenberg, Waters, Barnes and Tanenhaus 1984). Secondly, effects of phonological regularity on lower frequency words are eliminated in standard lexical decision tasks (pronounceable non-words). Phonological effects are only obtained in lexical decisions when the stimuli make it very difficult to discriminate words and non-words (Waters and Seidenberg 1985). Finally, words with common spelling patterns showed faster recognition latencies than words with uncommon spelling patterns. Unlike the effects of phonological regularity, these effects occur with both lexical decision and naming tasks (Seidenberg et al. 1984, Waters and Seidenberg 1985).

Such effects were also found in reading Chinese by adult native readers (Seidenberg 1985a). Seidenberg (1985a) performed a naming experiment comparing Chinese words which either did or did not contain orthographic cues to pronunciation. He found that the presence of orthographic cues facilitated pronunciation of low frequency words but had no detectable influence on high frequency words. This result was considered to be analogous to the frequency-by-regularity interaction in English. He concluded that high frequency words are recognized on a visual basis holistically even in Chinese logographic system. Based on such findings, Seidenberg and his associates (1985) proposed a time-course model, suggesting that skilled readers recognize a large pool of higher frequency words on a visual basis. Only lower frequency words, which are recognized more slowly, show phonological effects (letter-by-letter processing as in an alphabetic language).

A study by Chung, Chen and Leong (1988) focusing on young native readers of Chinese demonstrated similar findings that high frequency characters were processed faster than low frequency characters by all Grade 4, 5 and 6 readers. However, in processing pseudo characters, the reversed frequency effect was found from those Grade 4 subjects, especially among those less skilled Grade 4 readers who processed high frequency characters slower than low frequency characters. Chung, Chen and Leong (1988) attributed such findings to the deviation in both age and reading proficiency among three groups.

The above studies focus on the findings from first language word recognition either in English or in Chinese. What about second language readers? Do they demonstrate similar frequency effects to native readers? Meara (1984) found that among non-native readers, word frequency had the same effect upon second language readers of English readers, except the response time discrepancy between high frequency and low frequency words was considerably more marked than was the case for native English readers, where the recognition of high frequency words was slightly faster than low frequency words, though their overall performance was significantly faster than the non-native readers' (Meara and Morris, cited in Meara 1984). Other than Meara's study there is little research available, either on the frequency effects on word recognition processes of second language readers, or on word recognition processes in Chinese as a second language.

Orthographic complexity effects

In English, word length has often been equated with complexity owing to the "readability formulae" or the spelling-to-sound correspondences. The longer the word, the more complicated it is to learn. However, opposing views have also been noted. Bernhardt (1984) stated that "beginning L1 readers have fewer problems with longer words such as 'alligator' which is graphically unique and of high visual imagery than more graphically similar shorter and more frequent words such as 'them', 'they', 'their', 'this' and 'that' (p. 323)." Studies conducted with eye tracking instruments by Bernhardt (1986) had similar results with second language readers of German.

If the claim that word length equals word complexity remains a controversial issue in English word recognition, it may present less of a problem in Chinese. Since Chinese is a morphemic script, under the experimental conditions proposed by the present study, the problem of word length would not be salient. As has been discussed, each Chinese character occupies the same space and is the same length. There is no one character longer than another. Although, in actual reading text, many terms and words are polymorphemic in nature (for example: (zoo), for active or action; means subjects or animals; refers to yard or garden); nevertheless, the individual parts of such words are usually also independent words with independent meaning.

A more interesting question relates to the effects of orthographic complexity (namely, number of strokes). It may be useful to investigate whether or not stroke complexity has any marked effect on the word recognition processes of different readers as well as the interactions between word frequency and stroke complexity.

A few studies which focused on the orthographic complexity effect provided controversial findings. Kawai (1966) reported that in testing Japanese adults, fewer reading errors were made on Kanji (characters) with more strokes than on the fewer stroke ones. Kawai also controlled for word frequency, which had the predictable effect that high frequency words were recognized faster and more accurately than low frequency ones. Kawai argued that stroke complexity *per se* does not hinder the word recognition processes of proficient native character readers. His argument was supported by Taylor and Taylor (1983).

Yeh and Liu (1972), however, found that characters with more strokes were harder to process than the ones with fewer strokes. In their experiment, each of the test items consisted of two characters, both either complex or simple, and as participants tried to recognize character pairs, the complexity of the pairs gradually increased. Their findings indicated that the recognition threshold was longer for complex characters (15 or more strokes) than for simple ones (10 or fewer strokes).

Finally, Leong, Cheng and Mulcahy (1987) studied both the printed frequency and stroke complexity effects on the Chinese word recognition processes by skilled and less skilled readers of adult Chinese. They found that the high frequency words were processed significantly better (both faster and more accurately) than low frequency words, and the simple stroke (less than 10 strokes) characters were processed better than complex stroke (more than 11 strokes) characters by all subjects. However, it was the low frequency characters that contributed to much of the difference. The accuracy and latency deviation on stroke complexity effect and on frequency effect were more marked for less skilled readers than for skilled readers. For less skilled readers, the low frequency complex stroke characters were processed significantly more slowly and less accurately than high frequency counterparts. However, stroke effects did not affect their performance on high frequency characters. The results suggested that frequency effects and stroke effects influence the word recognition processes. The degree of interference depends on the reader's proficiency. The less skilled the reader, the more interference such effects cause.

Research into the effects of orthographic complexity on word recognition by native readers is, therefore, conflicting, and may indicate that complexity of characters *per se* is not necessarily the cause of difficulty in word recognition. As indicated by the review of literature, there is a need for more research on the impact of stroke complexity and printed frequency effects on second language word recognition processes in Chinese by readers with different reading proficiencies.

Thus, two research questions were investigated in the present study:

1. Does degree of familiarity with the words differentially affect the word recognition processes of L1 and L2 readers of Chinese? If so, what are the differences that L1 and L2 readers of Chinese demonstrate in their word recognition processes in terms of accuracy and response times?

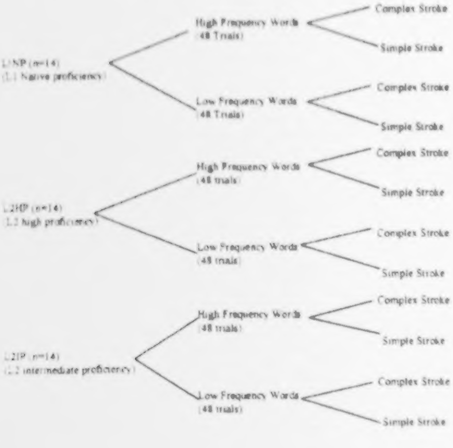
2. Does the complexity of the word structure differentially affect the word recognition processes of L1 and L2 readers of Chinese? If so, what are the differences that L1 and L2 readers of Chinese demonstrate in their word recognition processes in terms of accuracy and response times?

The Study

The research questions as to whether word frequency and stroke complexity have any impact upon word recognition processes in Chinese were explored in a context-free word recognition task on the computer involving three groups of adult readers selected according to L1 and L2 status and levels of Chinese reading proficiency. Group 1 consisted of 14 adult native Chinese readers (L1NP); Group 2 was made up of 14 non-native high proficiency L2 readers (L2HP); and Group 3 included 14 intermediate proficiency L2 readers (L2IP). All subjects completed a word recognition test and a reading comprehension test in Chinese to determine their levels of Chinese reading proficiency. The participants' background information is listed in Figure 1.

Figure 1

Research Design for Experiment 1: Context-free Word Recognition



The Method

The research examined the frequency and stroke complexity effects in Chinese word recognition processes:

1. by comparing L1 and L2 reading of Chinese to determine if there are any processing differences between the native and highly proficient L2 groups in word recognition processes;

2. by comparing proficient and less proficient L2 readers of Chinese to ascertain whether different levels of L2 reading proficiency have any impact upon the word recognition processes in second language.

All subjects completed a context-free word recognition task on computer.

Experimental Task

A same/different paradigm was used in the experiment. In this task, the subject saw a target Chinese word (along with three filler items) on a computer screen and then, after a brief interval, a test item was shown. The subject pressed one of two keys to indicate whether the test item was the same or different from the target word. There were equal numbers of *same* and *different* trials. For the 50% of trials on which the test item was, in fact, different, the test items (foils) were real Chinese words which resembled the target words either graphically, phonologically or semantically. Only one type of foil was tested on each *different* trial.

Materials for the Experiment

In this experiment only one-character words were used. The reasons for selecting only one-character words in the study was that, although one Chinese character is always one morpheme, a Chinese word is not always only one character. A Chinese word may be polymorphemic consisting of a combination of two or up to eight morpheme characters. Therefore, the length of the words may have some effect on the reading process, and the relation between each morpheme in a polymorphemic word may also add contextual facilitation or contextual inhibition effects on the reading process. These factors may affect the type(s) of encoding strategies that readers use in the word recognition process. Thus, in order to control all of these factors, which may have some bearing on the reading process, only one-character words were used in this study.

There were 192 trials in the experiment. Each trial involved the presentation of a display of four words (a target word and three fillers of the same frequency and complexity as the target). After the target presentation, a test item either the same or different from the target was displayed. In order to avoid a response bias, 50% of the trials were *same* trials.

Frequency/Stroke Considerations

All the words, including fillers, used in the experiment were selected according to printed frequency (high and low) from Cheng (1982), Liu, Chuang and Wang (1975) and Leong (1987) and orthographic structure (simple and complex as represented by the number of strokes). Since the number, order and directionality of strokes for a given character is invariant, stroke number was a good index of the internal structure of the Chinese grapheme (Cheng 1982). In this study a character which contained less than 9 strokes was considered to have simple structure. Any character that contained more than 12 strokes was considered complex. There were equal numbers of words in each frequency level and each stroke level. (Please see Figures 2 and 3 for details.)

Figure 2

Different Trials Involving Graphic Foil with HFCS (high frequency, complex strokes)

Visually Similar Pairs

pronounced "shu" meaning books

Target Display	書	輕	寫	福
Test Item (foil)	畫			

Figure 3

Different Trials Involving Graphic Foil with HFSS (high frequency, simple strokes)

Target Display	石	自	午	以
Test Item (foil)		目		

Note: a) In the actual display, no markers appeared;
b) The remaining 3 items in the display are fillers.

All the stimulus items and the response items, each measuring 1 cm square, were selected from a Chinese computer word card, programmed for the Apple Macintosh microcomputer system and shown centrally on the micro-computer screen for the experiment.

In summary there were 192 trials in this experiment, of which there were 96 *same* trials and 96 *different* trials. There were 576 characters serving as fillers in the target items with 3 fillers for each target item. The total number of characters used in the experiment was 768. All the characters were selected according to the printed word frequency in Chinese (high and low) and orthographic structure (simple and complex). There were equal numbers of characters in each frequency level and each stroke level.

Research Design for The Experimental Task

The experiment utilized a 3 x 2 x 2 x 2 factorial design with reader groups (L1NP, L2HP and L2IP) as a between-subjects factor and printed word frequency (2 levels: high, low), orthographic structure/stroke number (2 levels : simple, complex), and trial-type (2 types: same and different) as the within-subjects repeated measures.

The dependent measures were:

a) accuracy, that is, number of correct responses made by each subject on both *same* and *different* trials;

b) the response time measured in milli-seconds (msec), that is how fast the subject responded to each trial. The response time was measured from the beginning of the display on the screen until a key was pressed by the subject to indicate his/her judgment.

Planned comparisons were used on group contrasts to compare the L1NP with the L2HP groups and the L2HP with the L2IP groups. Since the group contrast was not orthogonal, the Scheffe confidence intervals were computed on group contrast to test if the differences between groups were significant. (The multi-variant approach to repeated measures was used.) The Wilks F-Test is reported. The analysis was conducted using MANOVA procedure (O'Brien and Kaiser 1985) in the 3.1 SPSSX statistical analysis package.

Procedure

A Macintosh MacPlus microcomputer was programmed to display the trial sets for the experimental tasks, to collect response accuracy and latency data, and to manage the randomization of the stimuli and the presentation order.

The subjects were tested individually in a quiet room. Each time, one target item and three fillers were shown on the screen for 400 milli-seconds with 20 milli-seconds interval between the target display and the test item. As soon as the target display disappeared from the screen, a test item, either a same or different item (if different, then one of the three foils) for the target was presented on the same spot as the target, to be judged either "same" or "different" to the stimulus item. The positions of trial items on the screen were equally randomized from position one to position four. The test item was shown on the same screen location as the target accordingly. The positions and the presentation sequences for both the stimulus items and the response items were counterbalanced within and across subjects.

4-9

Results and Discussion

1. Accuracy Measures on *Same* Trials

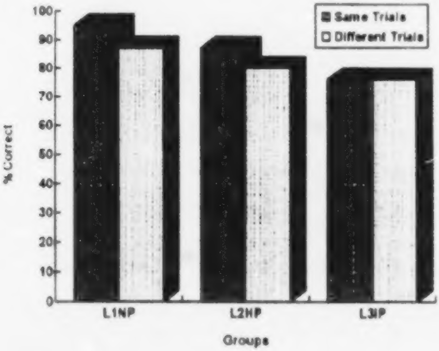
In this paper only the results on *same* trails were reported, to avoid the interference of additional information contained in the different trials. *Same* trials consisted of the presentation of a target item followed by a test item, which was identical and therefore warranted "same" judgement in order to be accurate. In contrast, *different* trials presented a test item which was different from the target item, and, as a result, required a "different" judgment in order to be accurate.

The following main effects were found to be statistically significant: group, frequency and stroke. In addition the group-by-frequency and group-by-stroke interactions were also found to be significant.

The group main effect, F(2, 39)=22.20, p<.0001, indicated that there were significant differences between the three groups in responding to *same* trials. Planned comparisons revealed significant differences between the L1NPs and the L2HPs, F(1, 39)=31.48, p<.0001; and between the L2HPs and the L2IPs, F(1, 39)=12.91, p<.001. The means for each group on *same* trials revealed that the L1NP group was most accurate (M=94.42, SD=3.65); the L2HP group was next (M=85.42, SD=6.64), and the L2IP group was the least accurate (M=74.85, SD=11.14). Figure 4 shows the three groups' performance on the *same* trials. The accuracy results on *different* trials are also presented in this figure in order to provide a comprehensive picture of the three groups' performance. As shown in Figure 4, the most proficient readers -the L1NP group- were most accurate in responding *same* when the test items were in fact the same as the target items. It is interesting to note that the accuracy rate decreased in linear progression as the readers' proficiency decreased and the differences in accuracy between *same* and *different* trial types also diminished as the level of reading proficiency decreased.

Figure 4

Group effect on the *same* and *different* trials

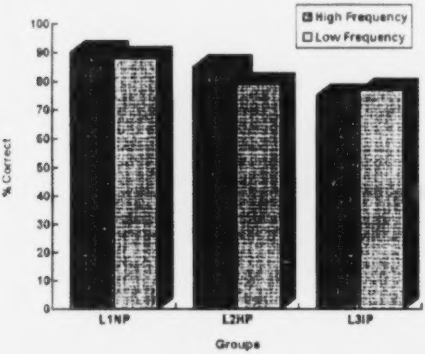


The frequency main effect, F(1,39)=18.96, p<.001, reflects the fact that accuracy performance on high frequency characters was higher (M=88.29, SD=10.30) than it was on low frequency characters (M=81.50, SD= 13.23).

The group-by-frequency interaction, F(2,39)=6.57, p<.005, is displayed in Figure 5. As indicated by Figure 5, the L1NP was most accurate on both high frequency and low frequency characters. Both L1NPs and L2HPs performed better on HiF characters than LoF characters. However, the discrepancy between accuracy rate on HiF and LoF characters was less obvious for the L1NPs than for the L2HPs. This could be attributable to the L1NPsÒ near 100% accuracy level which left little room for frequency effects to be obtained. The L2IPs' performance was different from both L1NPs and L2HPs. With the L2IP group, HiF characters were recognized even less accurately than LoF characters. The results may imply that HiF characters were still not familiar to the L2IPs. These observations were confirmed by planned comparisons which revealed that there was no group-by-frequency interaction effect between the L1NP and the L2HP performance, p>.10, but that there was a significant interaction between the L2HP and the L2IP, F(1, 26)=14.85, p<.001. The significant interaction effect was the result of the absence of the 'word familiarity' effect for the L2IP readers. The results are in line with some of the recent findings that sensitivity to word printed frequency is related to readers' reading proficiency (Leong, Cheng and Mulcahy 1987, Seidenberg 1985).

Figure 5

Group by frequency interaction



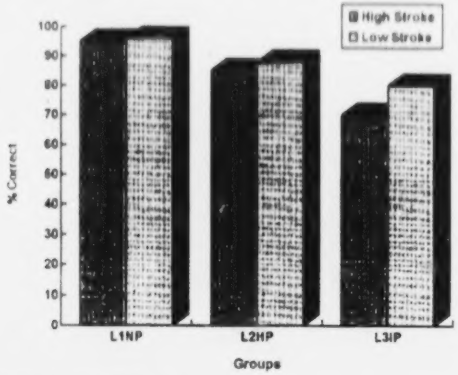
The stroke main effect, F(1,39)=20.22, p<.0001, showed that low stroke characters were processed more accurately (M=87.45, SD=9.92) than high stroke characters (M=82.34, SD=13.6). Thus, complexity of orthography plays a role in Chinese word recognition processes.

BEST COPY AVAILABLE

The significant group by stroke interaction, F(2,39)=8.69, p<.001 is shown in Figure 6. Planned comparisons indicated a significant difference between the L1NP and the L2HP groups, F(1,39)= 8.42, p<.005, and a significant difference between the L2HP and the L2IP groups, F(1,39)= 8.99, p<.005. As evidenced in the figure, native readers performed equally accurately (and near ceiling) on both complex (M=94.22, SD=4.92) and simple stroke characters (M=94.64, SD=5.40), while the least proficient reader group, the L2IP, performed more poorly on characters with complex strokes (M=69.05, SD=12.31) than simple strokes (M=80.65, SD=10.88). Thus, in this experiment, the more proficient the readers were, the less they were affected by the number of strokes in a character. This could indicate that the number of strokes in each character affects the word recognition process of less proficient readers more than that of proficient readers, but the ceiling effects in this interaction cannot be ruled out.

Figure 6

Group by stroke interaction for accuracy measures on *same* trials



2. RT Measures on *Same* Trials

The MANOVA analysis of general RT measures on the *same* trials yielded only the following significant main effects: group, frequency, and stroke. Only the group-by-frequency interaction was found to be significant.

The significant group effect, F(2, 39)=9.97, p<.0001), frequency effect (F(1,39)=15.85, p<.0001), and stroke effect (F(1,39)=40.06, p<.0001.) reflected essentially the same processing patterns as in the overall analysis:

a) The L1NP group once more was the fastest. Planned comparisons revealed a highly significant difference between L1NP and L2HP groups, F(1,39)=19.82, p<.0001, but no differences between the L2HPs and L2IPs, p>.10.

b) The frequency effect indicated that high frequency characters were processed faster (82.12 msec faster) than low frequency characters,

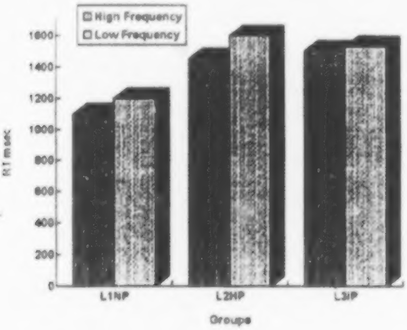
and

c) the stroke effect showed that simple stroke characters were processed 99.04 msec faster than the complex stroke characters in general.

The group-by-frequency interaction, F(2, 39)=5.14, p<.01 illustrated in Figure 7, again reflected similar patterns to the overall analysis. Both L1NP and L2HP readers demonstrated a similar degree of sensitivity to high and low frequency characters, with HiF characters processed much faster than LoF characters. Planned comparison further indicated a significant interaction between L2HPs and L2IPs on frequency, F(1,39)=10.26, p<.003. No significant difference was found between L1NPs and L2HPs, p>.10.

Figure 7

Group by frequency interaction



In summary, then, on *same* trials, the analyses of response time scores indicated that the processing times for both L2 reader groups were longer than for native, proficient readers.

In general, all three groups responded to high frequency and simple stroke characters faster than low frequency complex stroke ones. But the significant group-by-frequency interaction indicated that the least proficient group, L2IP, processed the frequency information differently from both the L1NP and the L2HP groups. The L2IPs were equally slow on both HiF and LoF characters. There was essentially no variation between the two frequency conditions for this group. With reference to their low accuracy rate on *same* trials, the RT results may imply that both HiF and LoF characters were equally unfamiliar to the least proficient readers.

General Discussion

It has been recognized that word frequency influences the word recognition process, while the effects of orthographic complexity still remain a controversial issue in Chinese word recognition. However, the present study clearly indicates that both word frequency and orthographic complexity influence the word recognition process in Chinese, especially for second language readers.

In line with Seidenberg's time-course model, the high frequency words in the present study were processed faster than the low frequency words. As indicated in the time-course model, whether recognition is direct or mediated depends on the time course of the decoding process. As salient orthographic units are recognized, they activate their phonological representations. In this way phonological access lags behind the visual analysis. 'Direct access' results when sufficient orthographic information is extracted from the input to permit recognition prior to access of phonology. This applies to a large pool of high frequency words. The slower recognition of lower frequency words allows more time for phonological information to accrue, either because sub-lexical orthographic patterns or the lexical items in a group of candidates activate their phonological representations. As a result, there will be phonological mediation only for the more slowly recognized lower frequency words.

The findings on the effect of orthographic complexity indicated that orthographic (stroke) complexity *per se* does not necessarily affect the word recognition processes of proficient L1 readers as shown in Figure 6. However, stroke complexity strongly influenced the L2 readers' word recognition, especially that of the least proficient L2IP readers. The results substantiate the findings from Kawai (1966), Leong et al. (1987) and Taylor and Taylor (1983), which focused on native character readers only; the proficient native readers' word recognition processes were not affected by orthographic complexity. At the same time the findings clearly indicated that, irrespective of whether word recognition was in L1 or L2, the influence of orthographic complexity depends upon the readers' target language reading proficiency. The less proficient the readers are, the more influence it will impose on their word recognition processes.

Finally, the findings on the effects of frequency and orthographic complexity provide us with more information about the time-course model. Each of the factors influencing processing, target language reading proficiency, orthographic complexity, and frequency, should be considered in using the time-course model to investigate the L1/L2 word recognition process in Chinese.

References

Bernhardt, E.B. (1991). *Reading development in a second language: theoretical, empirical, and classroom perspectives.* NJ.: Ablex.

Bernhardt, E.B. (1984). Toward an information processing perspective in foreign language reading. *The Modern language Journal*, 68:322-331.

Chitiri, H.F., Sun, Y., Willows, D.M. & Taylor, I. (1992). Word recognition in second language reading. In R.J. Harris (ed.). *Cognitive processing in bilinguals.* North-Holland, Amsterdam: Elsevier Science Publishers.

Chung, C.M., Cheng, S.C. & Leong, C.K. (1988). Interaction of reader and stimulus characteristics in children's recognition of Chinese. In I.M. Liu, H. C. Chen & M.J. Chen (eds.), *Cognitive aspects of the Chinese language.* Hong Kong: Asian Research Service.

Hoosain, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese.* Hillsdale, N.J.: Erlbaum.

Kawai, Y. (1966). Physical Complexity of the Chinese Letters and Learning to Read It. *Japanese Journal of Educational Psychology* 14:129-38.

Leong, C.K., Cheng, P.W. & Mulcahy, R. (1987). Automatic processing of morphemic orthography by mature readers. *Language & Speech.* 30:181-197.

Liu, I.M., Chuang, C.J. & Wang, S.C. (1975). *Frequency count of 40,000 Chinese words.* Taipei: Lucky Books.

Meara, P. (1984). Word recognition in foreign language. In A.K. Pugh & J.M. Ulijn (eds.), *Reading for professional purposes: strategies and practices in nature and foreign language.* London: Heineman.

O'Brien, R. & Kaiser, M.K. (1985). MANOVA method for analyzing repeated measures designs: An extensive primer. *Psychological Bulletin*, 97:316-333.

Seidenberg, M.S. (1989). Reading complex words. In G.N. Carlson & M.K. Tanenhaus (eds.), *Linguistic structure in language processing.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Seidenberg, M.S. (1985a). The time course of information activation and utilization in visual word recognition. In D. Besner, T.G. Waller & G.E. Mackinnon (eds.) *Reading research: Advances in theory and practice.* Vol. 5. Orlando, FL: Academic Press.

Seidenberg, M.S. (1985b). The time course of phonological code activation in two writing systems. *Cognition*, 19:1-30.

Seidenberg, M.S., Waters, G.S., Barnes, M.A. & Tamenhaus, M.K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior*, 23:383-404.

Taylor, I. & Taylor, M.M. (1983). *The Psychology of Reading.* New York: Academic Press.

Waters, G. & Seidenberg, M.S. (1985). Spelling-sound effects in reading: Time-course and decision criteria. *Memory & Cognition,* 13:557-572.

Yeh, J.S. & Liu, I.M. (1972). Factors Affecting Recognition Thresholds of Chinese Characters. *Acta Psychologica Taiwanica,* 14:113-117.

A WORD IN YOUR EAR: TO WHAT EXTENT DOES HEARING A NEW WORD HELP LEARNERS TO REMEMBER IT?

Monica Hill

One of the major differences between the English language needs of secondary and tertiary students in Hong Kong lies in the substantial vocabulary learning load faced by students at the start of an undergraduate degree course. A first year undergraduate faces the task of learning new subject specific vocabulary, as well as the general academic vocabulary which forms an essential part of academic discourse (cf. Laufer 1988 and 1992). An interesting question is how such students cope with absorbing this vast new quantity of words. Many rely on guessing the meanings from context, while others tend to check the meanings in either English-Cantonese dictionaries or, in a few cases, in monolingual dictionaries (Bensoussan and Laufer 1984). Once they have learned, or partially learned, a new word, students often try to use it again either in their written assignments, or in discussions and oral presentations in class, with varying degrees of success.

A related problem is that students are often unable to reproduce a 'new' word orally as they are unsure of its pronunciation. Some students are able to read the phonemic transcription of a new word in its dictionary entry, and reproduce it with reasonable accuracy. Many others, however, either do not understand the system of phonemic transcription, or do not bother to look at the transcription (Piczon-Llamzon 1979). These students in particular have a great deal of difficulty with oral reproduction of new words and are particularly prone to errors in word stress. Irregular pronunciation such as / krɪ'eɪtʃə / for creature, / ekstra'veɪgənt / for extravagant / 'meɪtʃə / for mature and /monəu'geɪmɪ/ for monogamy illustrate some of the phonological difficulties which Hong Kong tertiary students face with partially learned words. (McNeill 1990). This suggests that the phonological dimension of word learning is not, or is only partially covered. In the Hong Kong education system, relatively little attention is paid to the teaching of pronunciation, which means that such errors are widespread and pass largely undetected. Obviously, this lack of phonological awareness on the part of ESL students is not really apparent in their written work, where they are generally able to produce correct word forms. Lexical errors in writing tend to be semantic in nature or involve incorrect collocation.

An additional factor which needs to be taken into account in an investigation of Cantonese-speaking students' difficulties with English vocabulary is the effect of word length. Given that Chinese contains many one character words, a reasonable hypothesis might be that native speakers of Cantonese find it more difficult to remember and reproduce English words of more than one syllable. In fact, Chinese words do not necessarily consist of only one character and one morpheme: they can consist of a combination of polymorphemic characters which range from two to eight morpheme characters in length. As each of these polymorphemic Chinese words is made up of a series of semantic clues, learners may have more contextual information, or inhibition, to help them learn such words. (See Sun, this volume.) In the present study, comparisons are made to ascertain whether or not students encounter more problems over learning longer English words rather than shorter ones.

The purpose of this paper is to report the results of a study carried out at the University of Hong Kong which tried to identify the effects of phonological processing on first year students attending English enhancement courses in the English Centre. The first part (Preliminary Study) aims to identify a set of academic vocabulary items which are unfamiliar to the population from which the sample of subjects is drawn. The second part (Main Study) is experimental and investigates the learning of new vocabulary under different conditions.

Research Questions

The study addresses two questions:

1. Do learners who can read the phonemic transcription in dictionary entries have any advantage in their learning of new vocabulary?

2. Do learners who hear new words as well as see them have an advantage over learners who only have access to the written form?

Hypotheses

In attempting to answer the above questions the following three hypotheses are tested:

1. Phonological processing * of new vocabulary improves learners' overall ability to learn new words.

2. Phonological processing at the initial learning stage improves learners' ability to produce new words in writing and in speech.

3. In view of the polymorphemic structure of Chinese words, long English words will be more difficult to learn than short words. **

* For the purpose of the study, phonological processing includes either the hearing of the new word or the phonological decoding of a written word form using phonemic transcription.

** For the purpose of the study a long word is defined as a word which consists of three or more syllables and a short word is defined as consisting of one or two syllables.

Design of Study

The study is divided into two phases: The aim of the preliminary study was to identify a set of academic words which were largely unknown to the subjects, while the main study set out to examine the effect of phonological processing on Cantonese-speaking learners of new English words.

Preliminary Study: Identifying a set of target words

Based on a University Word List (Nation 1990), a list of approximately 250 words was circulated to a target sample of 58 students from the Arts and Social Sciences Faculties. They were asked to delete any words whose meanings they knew and to mark those words which they had not previously encountered. (Words were considered 'known' if a student recognised the form and could reproduce the meaning in his/her own language.) The students returned all the papers and were advised not to try to memorize any of the words or look them up in dictionaries as they would be given the word lists again at a later date, if they so wished. The results of the survey varied considerably, and only five words were claimed to be completely unknown in the subjects' self reports. Subsequent objective testing showed that their claims of word knowledge had been exaggerated. The 45 least frequently known words were extracted and isolated to form the basis of the pre-test. (See Appendix 1)

The group of words was further revised according to syllabic length and produced a list of 30 words in the following patterns:

1. 7 words of 1 syllable
2. 8 words of 2 syllables
3. 7 words of 3 syllables
4. 8 words of 4 or more syllables.

Main Study

To further ensure that the words used in the experiment were actually new to the subjects, a pre-test was used. Each subject was given a sheet showing the original 45 least frequently known words and was asked, if possible, to write the L1 translation or give an L2 equivalent. If they did not know the word, they were asked to leave a blank.

They were asked to give their names, class and 'Use of English' A level result, where appropriate, so that a closer analysis of the findings might be made. In this way, the subjects were not immediately shown the 30 target words, but were gradually introduced to the level of words to be learned.

Study Phase

Three dimensions of word knowledge, identified by Aitchison (1987) as semantic, syntactic and phonological, have all been reflected in the experiment and are described in Figure 1.

Figure 1

Elements of the Presentation Materials

Dimension of word knowledge	**Operationalisation in presentation materials**
semantic syntactic phonological	L1 translation use in context phonemic transciption and aural element on tape

The words were first listed in alphabetical order. To replicate the methods currently used by the students to learn new words, the lists showed the following:

Figure 2

Example of New Word Explanation Used in Study Phase

Word in English (L2) phonemic transcription	**(Chinese) L1 translation**	**Monolingual (L2) dictionary explanation (i.e. word in context)**
auspices / ˈɔspisiz /	贊助	The conference was held under the auspices of the United Nations.

The Chinese translation was taken from *Longman's Active Study English - Chinese Dictionary (1991)* and the monolingual explanation (i.e. contextualised meanings) from *Collins Cobuild Dictionary*. An additional ingredient was added to give more authenticity to the exercise of replicating the students' dictionary search: namely the phonemic script, as shown in Longman's Active Study Dictionary. This also helped to balance out the phonological element of the other half of the experiment. In this way, the students were able to see the target words and their pronunciation, have the translation in L1 and the word used in context in L2. The Chinese translations were kindly transcribed by Mr K.K. Chan of the Language Centre, University of Hong Kong.

The dictionary explanations chosen generally reflected the most academic meanings, for example, 'gem' as in '... a gem of wisdom...' rather than 'a precious stone'. Slight changes were made, for example, to alter plural nouns to singular ones, as in 'these anomalies' - 'this anomaly' to remove unnecessary obstacles to the subsequent reproduction of the target word.

The students were given fifteen minutes to study the 30 words which had been selected according to word length before being asked to reproduce them both in writing and orally. To enhance the phonological dimension of word learning, the experimental group was given the same word study sheets but their experiment was conducted in the language laboratory and they were able to hear the target words pronounced on tape. They were allowed to listen to the tape as often as they wished and were able to repeat the words silently or aloud as they wished. The time limit of fifteen minutes was the same for all subjects.

Productive Phase

The target words were listed in random order so that students were not able to rely on recalling the words in the same order as originally presented. In the random arrangement, the lengths of words as well as their initial letters were varied. The subjects were given the contextualised meaning with a gap for the target word, plus the Chinese translation and were required to produce the target word. They were asked to give any words, or any parts of the words they could remember and were advised that immediately after the written test, they would be asked to record on tape the target words, or the parts that they could remember. The aim of this was to distinguish more accurately between the effects of aural and visual dimensions of pronunciation.

To ascertain the extent of the subjects' knowledge of phonemic script, they were given a short questionnaire. The purpose here was to help identify those who used their previous knowledge of phonemic transcription and those who disregarded or were unable to use this information in the test.

Both groups were asked to repeat the recall phase a week later. This helped identify the words which were remembered in both mid- and long-term memory (referred to here as tests 1 and 2).

General English standard of the subjects

A total of 56 students from the Arts and Social Sciences faculties took part in the experiment, with 28 in each group. By calculating the mean Use of English grades of the subjects, both the control and experimental groups were evenly balanced at C6. The lowest level accepted for admission to the University of Hong Kong is D8, with A1 as the highest grade on the scale, so these were fairly average first year tertiary students.

Results and Discussion

Results of the Pre-test

Both the experimental and control groups were fairly evenly balanced in terms of their familiarity with the target words. The mean number of known words was 1.0 in the experimental group and 1.7 in the control group. In many cases, however, the correct answer in the pre-test appears to be an inspired guess as the wrong answer was produced in the subsequent tests. For this reason, the apparent number of known words is noted in the data, but not deducted from the final scores as it does not necessarily reflect an accurate picture of the subject's familiarity with the word.

Results of Main Study

The answers to both the short-term and the long-term tests were entered on to a database (FileMaker Pro) and evaluated. Spelling mistakes counted as errors, but third person 's' or past tense forms e.g. 'pervades' or 'pervaded' for 'pervade' were accepted if they were contextually and grammatically appropriate, while 'auspice' was incorrect as only the plural form 'auspices' is used.

In general, there was no significant difference between the two groups in the written form mid-term recall (Table 1). However, the difference between the mid-term and long-term recall produced a decrease of 27.6% for the experimental group and 32.3% for the control group. The difference in the scores of the Control and Experimental groups on Test 2 (long-term retention) is statistically significant ($t = 2.33$, $p < .05$). The addition of the sound dimension, therefore, appears to have been an advantage for the experimental group in the overall long-term recall.

Written recall

Figure 3

Written Recall Scores

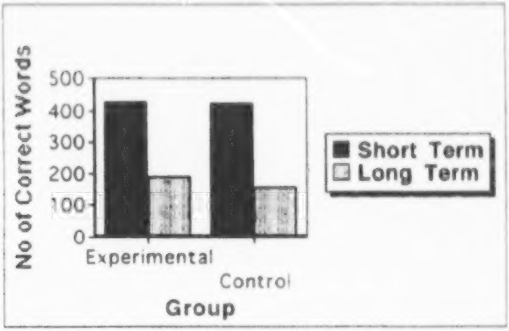


Table 1

Written Recall Scores

	Experimental (n = 28)		Control (n = 28)		Notes
	mean (/840)	%	mean (/840)	%	
Test 1 (mid-term)	428	50.1	424	50.5	(t = n.s.)
Test 2 (long term)	189	22.5	153	18.2	(t = 2.33, p<.05)

Spelling Errors

The control group generally produced more accurate spelling as their words seemed to have been produced from visual memory, resulting in fairly high scores on the written tests. The experimental groups appear to suffer from less accurate spelling with common errors being:

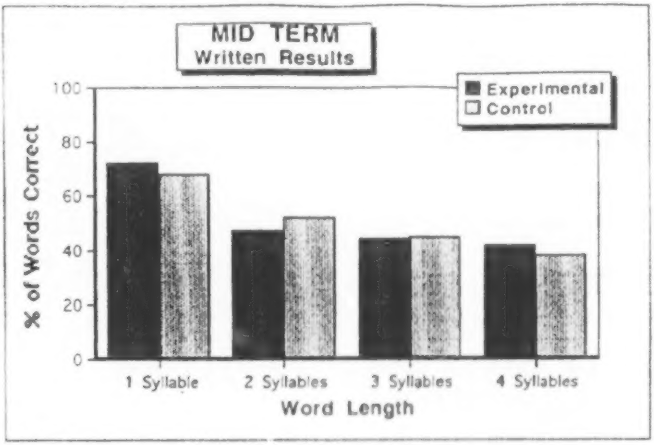
'superci**ll**ious', 'amo**phr**ous', 'rever**br**ate' and 'p**re**vade'

Perhaps this accuracy of spelling in the control groups reflects the Chinese skills of character recognition and memory, as, in initial learning of Chinese characters, they are taught to recognise them visually at school. Those in the control groups tended to write down the words on their sheets as if writing them helped to recall them visually while those in the experimental groups tended to repeat the words quietly to themselves. The additional input of sound in the experimental group seemed to cause them a degree of difficulty, perhaps because the greater input of information was more difficult to 'digest', resulting in less accurate reproduction of the target words. Another explanation of the loss of written accuracy in the Experimental group is that the aural mode of presentation was unfamiliar to them and may have interfered with their usual word learning processes.

Closer analysis of the mid-term written results reveals the following:

Figure 4

Test 1 (mid-term)
Written Recall By Word Length



The hypothesis that Chinese-speaking learners would find long words more difficult to learn than short words was confirmed: 180 long words compared to 248 short words were recalled by the experimental group and 174 long words compared to 250 short words by the control group (Table 2).

Table 2

Test 1 (mid term)
Written Recall by Word Length

	Experimental (n = 28)		Control (n = 28)		Notes
	mean (*/196 **/224)	%	mean (*/196 **/224)	%	
1 syllable*	142	72.6	133	76.9 }	(t = n.s.)
2 syllables**	106	47.3	117	52.2 }	
3 syllables*	87	44.3	89	45.4 }	
4 syllables**	93	41.5	85	37.9 }	

There was no significant difference between the experimental and control groups when the data was analyzed by word length, although the experimental group did slightly better on words of 1 and 4 syllables while the control group took a slight lead in words of 2 and 3 syllables. Surprisingly, the experimental group found 3 syllable words the most difficult to recall in the mid-term.

In the long-term, however, this pattern changed slightly.

Figure 5

Test 2 (long term)
Written Recall By Word Length

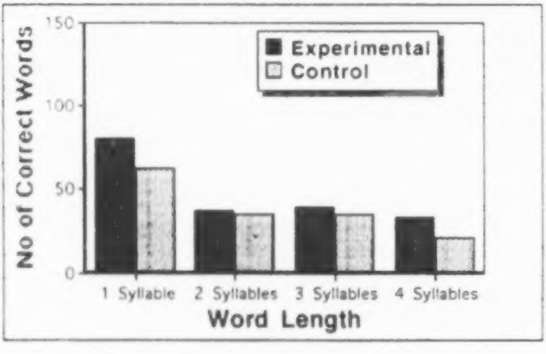


Table 3

Test 2 (long term)
Written Recall By Word Length

	Experimental (n = 28)		Control (n = 28)		Notes
	mean (*/196 **/224)	%	mean (*/196 **/224)	%	
1 syllable*	80	40.8	62	31.6	n.s.
2 syllables**	37	16.5	35	15.6	n.s.
3 syllables*	39	19.9	35	17.9	n.s.
4 syllables**	33	14.7	21	9.4	(t = 2.8, p<.05)

The experimental group outperformed the control group in all categories of word length. Short words (1 syllable) were recalled twice as frequently as the longer ones. With the words of 4 syllables, however, in long-term recall, t-tests showed a significant difference with a t-value of 2.18 ($p<0.05$).

Oral recall

In this area the difference between the two groups' results proved significant.

Figure 6

Oral Recall Scores

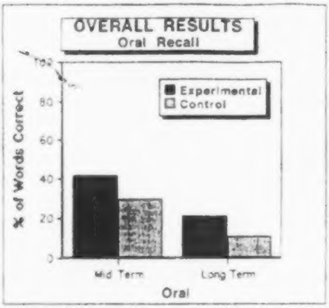


Table 4

Oral Recall Scores

	Experimental (n = 28)		**Control (n = 28)**		**Notes**
	mean (/840)	**%**	**mean (/840)**	**%**	
Test 1 (mid-term)	354	42.1	250	29.8	(t=2.81, p<.05)
Test 2 (long term)	178	21.2	96	11.4	(t=4.25, p<.01)

The variation between mid-term and long-term recall was a decrease of 50% for the experimental groups and 38.4% for the control groups, which was slightly less than the drop for written recall. A comparison of the two groups' results in Test 2 (Oral Recall) reveals a highly significant difference (t=10.92, p<.01).

The experimental group produced 70% more correct words in the mid-term and 54% more in the long-term, again confirming that the phonological dimension improves learners' overall ability to learn new words.

Closer analysis of the oral results shows the following:

Figure 7

Test 1 (mid term)
Oral Recall By Word Length

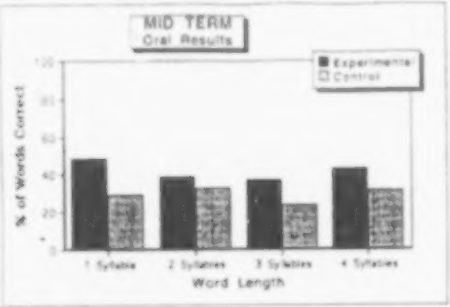


Table 5

Test 1 (mid term)
Oral Recall By Word Length

	Experimental (n = 28)		Control (n = 28)		Notes
	mean (*/196 **/224)	%	mean (*/196 **/224)	%	
1 syllable*	96	49.0	57	29.1	t=3.92, p<01
2 syllables**	88	39.3	74	33.0	n.s.
3 syllables*	73	37.2	47	24.0	t=2.34, p<.05
4 syllables**	97	43.3	72	32.1	t=2.03, p<.05

Contrary to the written results, the experimental group found words of 1 and 4 syllables easier to recall orally than those of 2 and 3 syllables (Table 5). The control group had significantly more difficulty in oral recall in all word lengths, not having heard the words pronounced on tape, and/or ignoring the phonemic transcription given on the sheets.

These subjects often had no idea how to say words such as 'beige' and 'niche', perhaps as the words looked unusual as well as unfamiliar. Some students simply omitted to 'read' these words on to tape, completely ignoring the answers they had written on their sheets, while others said, 'Don't know how to pronounce!' or spelled out the words, e.g. 'b-e-i-g-e'.

An interesting phenomenon which emerged in the subjects' oral production of longer words was the tendency to add an extra syllable to the word, e.g. /ə mɔ'fəs/ frequently became / ə mɔ'farəs /. Although the longer 4 syllable words were not so easily recalled overall, they did not pose such a relatively great problem in spoken form.

Long-term oral results

Here the effect of the phonological processing on the experimental group was clearly discernable. Those who had heard the pronunciation of the words did significantly better in the oral recall than those who had not. Again the disparity was clearest in words of 1 syllable, as those who had not received aural clues were unable to pronounce correctly short words like 'gem', 'chore', 'beige', 'niche' or 'fraud'. These were most commonly recorded as /gem /, / kɔː /, / bɪdʒ /, / nɪtʃ/ and / frɔːd /

Closer analysis shows the following:

Figure 8

Test 1 (long term)
Oral Recall By Word Length

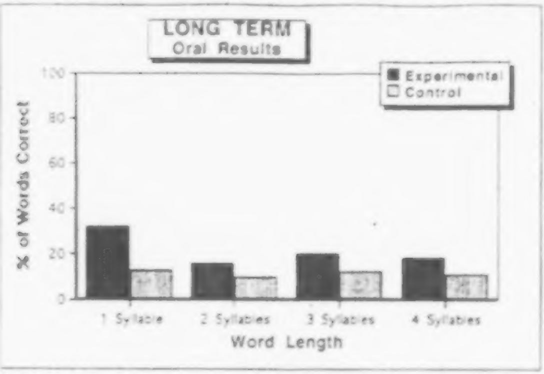


Table 6

Test 1 (mid term)
Oral Recall By Word Length

	Experimental (n = 28)		Control (n = 28)		Notes
	mean (*/196 **/224)	%	mean (*/196 **/224)	%	
1 syllable*	63	32.1	26	13.3	t=4.24, p<.0
2 syllables**	36	16.1	22	9.8	t=2.97, p<.05
3 syllables*	39	19.9	24	12.2	t=2.65, p<01
4 syllables**	40	17.9	24	10.7	t=2.64, p<01

The experimental group recalled short words more easily than long words, while there was no significance in the length of words recalled by the control group. Given the very low figures involved with the latter group, this is probably not a significant feature.

Conclusions

Although the number of words recalled in the long-term is rather low, the experiment has been able to test the effect of the phonological dimension on word learning and has shown that the sound dimension has considerable impact on learning new words, particularly on oral recall.

It has also shown that Chinese-speaking students do generally have more difficulty in learning longer words, although they often compensate by giving such words more attention, particularly in their written form. Longer words do not, however, pose such a problem in oral reproduction as they can often be broken down into smaller component parts and can involve fewer problems with stress.

Recommendations

Follow up work should examine the molecular breakdown of the syllables recalled. This may show which parts of the words are most frequently recalled, e.g. initial sound / medial / or final sound and in which combinations.

The results suggest that further studies should be considered to assess the importance of:

1. overt teaching of phonemic transcription on vocabulary learning from dictionaries and its effects on the pronunciation of words

2. increased phonological processing, such as ensuring that all subjects pronounce the words as they are learning them

3. 'talking' dictionaries and computer programmes which have 'real sound' as opposed to digitized voice reproduction, and their effect on vocabulary learning.

References

Aitchison, J. (1987). *Words in the Mind*, Oxford: Blackwell.

Bensoussan, M. & Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading* 7(1):15-32.

Carter, R. & McCarthy, M. (eds.) (1988). *Vocabulary and Language Teaching*. Harlow: Longman

Laufer, B. (1988). What percentage of text-lexis is essential for Comprehension? In Chauren & M. Nordman (eds.) *Special Language*, Cleveden: Multilingual Matters 1988.

Laufer, B. (1992). Reading in a foreign language: how does L2 lexical knowledge interact with the reader's general academic ability? *Journal of Research in Reading,* 15(2):95-10.

McNeill, A. (1990). Vocabulary learning and teaching: evidence from lexical errors in the spontaneous speech of ESL learners. *Institute of Language in Education Journal* 7:141-153.

Nation, I.S.P. (1992). *Teaching and Learning Vocabulary.* New York: Newbury House.

Piczon-Llamzon, P. (1979). Reading the dictionary for the pronunciation of English words. *Guidelines* 2:49-53.

Rodgers, T.S. (1969). On measuring vocabulary difficulty: an analysis of item variables in learning Russian-English vocabulary pairs. *IRAL* 7:327-343.

Webb, W.B. (1962). The effects of prolonged learning on learning. *Journal of Verbal Learning and Verbal Behaviour* 1:173-182.

Xue Guoyi & Nation, I.S.P. (1984). A University Word List. In I.S.P. Nation (1992). *Teaching and Learning Vocabulary.* New York: Newbury House.

Appendix 1

Wordlist used in the study

allude
amorphous
anomaly
ascribe*
assent*
auspices
averse*
beige
chore
cogent
confer*
construe
contingent*
cumbersome*
detriment*
discern*
emaciate
fraud
gem
hunch
incessant*
ingenious*
ingenuous
interlude*
kindred
litigate
niche
partisan
perpetrate
pertinent
pervade
procure
propensity
purport
recur*
repudiate
reverberate
rudimentary
squint
stipulate*
suffice
supercilious
unduly
upsurge*
wrench*

All these words were used in the pre-test, in order to identify which, if any, were known to the students.

*These words were not used in the main study.

PIG IN THE MIDDLE? EFFECTS OF MEDIATING TASKS ON COGNITIVE PROCESSING OF TEXT.

Desmond Allison, Vivien Berry and Jo Lewkowicz

Introduction

The notion that exercises or activities may mediate between a text and student readers[1] appears to need little comment in itself, as this is clearly what they are designed to do in order to assist understanding. The exact nature and impact of such mediation, however, is more controversial, and raises questions over the most appropriate uses of class time and the proper roles of teachers and learners. A thorough examination of these concerns calls for a combination of reflection on experience and empirical investigation across a range of learners, activities and texts. The authors view these approaches to research not as mutually exclusive or opposed, but as complementary and indeed symbiotic.

This paper reports on a study of the effects of a written reading comprehension exercise and an oral discussion activity as alternative tasks mediating between a written text and learners' summaries of that text. Summary writing was considered to be a relevant activity for the learners in the study, who were first-year undergraduate students of economics. Summary writing, a task required by subject lecturers, calls upon abilities to comprehend texts, extract essential information and ideas, and write clearly and accurately (all points made by economics staff in discussions when the English course for first-year students was being designed). In a language class, oral or written summaries can also help teachers discern problem areas in students' comprehension (Carrell and Eisterhold 1988:88).

The study compares the two mediating tasks with each other and with a third condition in which the summaries were written with no mediating task. It must also be pointed out that summary writing may itself influence the interaction between text and readers, and that a summary is not a direct embodiment of a learner's interpretation of the text. In presenting and discussing the study's empirical findings, the authors do not pretend to resolve or even treat all the complex evaluative issues that arise. We do, however, seek to set some limits on speculative enquiry by relating possibilities to observed outcomes.

Stubbs (1987:27) points to the hidden curriculum that is constructed and enacted through choices of classroom activity, and affirms that "students' interpretations of texts can be widely different under different classroom practices." A reason for concern over such induced readings is that widely differing interpretations of texts can all be valid (Alderson and Urquhart 1984:47). An acceptance of plurality, and a corresponding rejection of any belief in one sole correct interpretation (the teacher's) of what a text can mean, do not of course imply that all interpretations are automatically valid or equal in worth, nor need these insights invalidate classroom practices that are aimed at developing learners' comprehension. Even so, the recognition that classroom practices will also lead to different readings of texts poses a challenge to the teacher When are we leading learners towards a better understanding of a text, and towards more effective strategies for coming to grips with other texts, and when are we prescriptively or unwittingly placing obstacles in the path of learners as they seek to come to terms with texts in their own diverse ways? When one also considers the complex relationship between understanding a particular text and integrating one's reading within personal content schemata in some domain of experience or study, the role of comprehension exercises or activities in the language class is no longer something that we can afford to take for granted.[2]

Learners' interpretations of texts are not directly accessible, but have to be inferred from other evidence. Many integrated tasks in language classes culminate in sustained language production by learners, and the resulting products offer indirect but important testimony for an account of how texts have been interpreted. With tertiary students this evidence might take the form of a written report, summary or essay, a spoken presentation on the topic, or a seminar-style discussion of issues raised in a text. Our focus here will be on final written outcomes.

Working with secondary school teachers as advanced learners in Hong Kong, and concentrating on how these learners reacted to different forms of reading comprehension exercises, Allison (1989) found that reading exercises that had been specifically designed for particular texts were substantially better received than were standard reading exercises of the kinds promulgated by Scott et al. (1984), Edge (1985) and Walker (1987). Allison recommended that further research could more usefully compare text-specific reading exercises with other activities, especially discussion of the text, in order to determine the nature and effects of these different mediation processes. In a first language context, research into the psychological processes involved in student writing already suggests that, under certain conditions, students' writing will be qualitatively different depending on whether the stages gone through before reaching the final written draft involve oral or written processes (Jenson and DiTiberio 1984).

Aims and Objectives of the Research

In the light of the above interests and concerns, a study was undertaken with first-year undergraduate students at the University of Hong Kong, the purpose of which was to determine how, and to what extent, students' written summaries of a text would differ under three different conditions. These comprised an oral mediating task (small group discussions), a written mediating task (completion of tailor-made reading comprehension exercises), and no mediating task (control group) between reading text and written summary.

Specific attention was paid to the following parameters:

- impression grades given by two independent raters (teachers not familiar with the research design);
- length and summary in terms of number of words;
- number of t-units (a t-unit being a main clause with any subordinate clauses: Hunt 1965);
- number of error-free t-units;
- categorisation of t-units (and also of error-free t-units) in terms of text dependence/ independence;
- inclusion of idea units 'targeted' by the researchers as relevant and reasonably likely to occur in a summary of the selected text.

Method

The 80 students in the study were pseudo-randomly assigned to one of three treatments[3]:

1. Group 1 (N = 28): reading text + oral mediating task + written summary
2. Group 2 (N = 26): reading text + written mediating task + written summary
3. Group 3 (N = 26): reading text + no mediating task + written summary (control group)

Comparison of the students' H.K.E.A. Use of English examination results, which afforded a recent and readily available independent measure of English language proficiency, revealed no significant differences between the groups. This suggests that between-group comparisons can reasonably be associated with the independent variable in our study rather than any accidental anomaly in the composition of the groups.

Each treatment group was given the same reading text, entitled "Is there a gene for genius?" (McCrone 1993). The text (See Appendix 1) of 1736 words, with a FOG index of 14.29 (Gunning 1952, Davies and Irvine 1993) was considered to be of appropriate reading difficulty level for the target population. The topic was judged to be of likely interest to these students, some of whose first-year courses were taken in common with students reading for a degree in Social Sciences. The article was taken from a quality newspaper. While this means that the text has elements of journalistic style rather than an academic textbook style, it clearly fell within the kind of reading material prescribed or encouraged in undergraduate degree curricula.

After reading the text, students in Group 1 discussed the text in small groups before writing their summaries (the discussions were recorded and will be analyzed at a later stage); students assigned to Group 2 completed a series of tailor-made reading comprehension exercises before writing (see Appendix 2 for details of this task); students in Group 3, the control group, were simply asked to read the text and summarize it in writing. All three groups were given a double class to complete the set activities. This was judged appropriate on the basis of a pilot study conducted on a comparable group of students. For groups 1 and 2 the time was divided into 20 minutes of reading time, 20 minutes on the mediating task and 40 minutes for writing the summary. Group 3 students were given the full 80 minutes to use as they considered most appropriate. The final task read as follows:

> On the paper provided,
>
> - summarize the article
> - evaluate the two main viewpoints developed in the article
>
> You may use the text and your notes (if any) to help you complete this task.

The last comment only was varied for Group 2 in light of their task, and read:

> You may use the text and the reading questions to help you complete this task.

Both the text and any other written materials were therefore retained while students wrote their summaries.

Each written summary was independently graded by two markers (in addition to the researchers) across the three conditions. In all there were six independent markers, all native speakers of English; three were male and three were female. Each marker graded 26 or 27 scripts; each script was marked twice by one male and one female marker; scripts were pseudo-randomly distributed to markers so each set of scripts represented all three conditions. The scripts were graded holistically on a 9-point scale using course grading guidelines that were familiar to all the markers. The scripts were also scored by the researchers on a number of quantitative and qualitative measures that could provide other bases for comparison across the groups. Word counts used the operational definition of "words" that derived from the insertion of spaces between visible "words" in the written texts (which were typed up for later ease of reference). Division of each text into t-units and subsequent identification of "correct" t-units in terms of standard grammar, spelling and vocabulary usage were tasks independently undertaken by each researcher. Any initial discrepancies in analysis were carefully resolved thorough discussion. The main aim here was to ensure consistency of judgement, so that meaningful comparisons could be made across the groups.[4]

Perhaps more controversially, but within the same operational spirit and context, each t-unit was also placed in one of four categories to denote the degree and kind of text dependence or independence that it exhibited. This approach owes much to Campbell (1990), but the categorisation was adapted for our work, not least as a summary writing task constrains what is to be expected. Direct attributed citation (which proved to be extremely rare) was included along with other "exact copies". In cases where the summary text differed from the source text, the concern was to describe the degrees of difference, for example in the extent of grammatical reconstruction or lexical substitution involved when "chunks" of source text appeared to have been adapted or combined. Any imposition of categories on such complex and graduated forms of differentiation will inevitably set up "boundary problems" for the analysts. We accepted these intellectual and procedural constraints in the interests of obtaining comparable data over a relatively large number of scripts, while naturally recognising that the richness of each individual script cannot be wholly captured in such descriptions. (We see in-depth analysis of individual scripts, protocol analysis and interview data as valuable research techniques that can complement the comparisons across groups featured in the present study.) This approach gave us the following categories:

decreasing in text dependency ↓

- exact copy = copied word for word correctly from the text
- near copy = small elements of the text added or otherwise exact copies; blend of extensive exact copying plus minor elements of paraphrase
- paraphrase = substantial syntactic reworking; bringing textual elements together
- own words = bringing ideas together or adding to the text.

The last three categories were identified in terms of the dominance of almost exact copying of source text, extensive grammatical reworking or recombination that still remained close to identified chunks of source text, or extensive departure from wordings in the source text. An example of each category (with intervening t-unit boundaries shown by the symbol //) is given below:

Exact copy:	The discovery of a first gene does not mean that the riddle of intelligence has been solved (script 70:8)
Near copy:	However, the children who fared best were those parents were both supportive and stimulating (script 55:18)
Paraphrase:	According to the result of Dr. Robert Plomin's experiments, an unnamed gene which plays a part in determining intelligence has been identified by using new gene mapping techniques (script 37:4)
Own wording:	Remember we have fertilizer in the world, // although the soil is not rich, fertilizer can improve it. // Parental influence is somewhat a kind of fertilizer for the growing up of a genius. (script 21:12,13,14)

The final measure for this study was the occurrence of identified "idea units" in summaries of the text (Appendix 3). It would have been unrealistic and excessively prescriptive to seek to list all and only those propositions that would occur in some ideal summary, and no such ambition informed our work. Our aim, rather, was to list selected units that the researchers considered to be relevant and reasonably likely to occur in summaries of this text. Our basis for such an analysis was:

1. independent listing of units by each researcher;
2. comparison with a summary written by a teacher;
3. several readings of the students' summaries (without specifically targeting this issue). The distribution of these "targeted" idea units could then be traced across the summaries written under the three task conditions.

The number of targeted points was compared across conditions. Any notable differences found across the conditions must remain open to qualitative interpretation. Further study of individual "idea units" will be undertaken in a later research report.

Evaluative interpretation of this sort of data is, of course, essential. While it is true that the identification of "idea units" in differently worded summaries already involves a degree of interpretation - and of intersubjective verification - this interpretative process must be sharply distinguished from any value judgements that might be made about the scope or effectiveness of different summaries. In particular, "successful" summaries could be written at various levels of detail, so that no simple relation between occurrence of points and quality of summary should be assumed.

that no simple relation between occurrence of points and quality of summary should be assumed.

Results and Discussion

Table 1 presents summary statistics for scores awarded in each condition. The results show that the control group (group 3: no intervening task) achieved higher average scores than group 2 (written mediating task). Group 2 in turn scored higher on average than group 1 (oral mediating task).

Table 1a

Mean Scores for Markers Across Three Conditions

Group	Count	Mean (max=18)	SD	F-ratio	F-prob
Grp 1	28	6.107	2.699	2.628	.079
Grp 2	26	3.470	3.470		
Grp 3	26	3.091	3.091		
Total	80	3.156	3.156		

Table 1b

Mean Scores for Markers + Researchers Across Three Conditions

Group	Count	Mean (max=27)	SD	F-ratio	F-prob
Grp 1	28	10.286	4.162	1.573	.214
Grp 2	26	11.192	4.561		
Grp 3	26	12.308	3.813		
Total	80	11.2375	4.219		

Analysis of variance (one-way Anova, SPSS for Windows: Norusis 1992) in Table 1a indicates that the difference in mean scores is not significant at the <.05 level. The observed tendency towards difference (p = .079) should not, however, be automatically dismissed. Further research with a larger sample size is needed to pursue this already noteworthy trend.

Observed differences are smaller when the researchers' marks are added to those of the markers (Table 1b). It is likely that the researchers had different expectations from the other markers. They were, after all, more familiar with the task and may have been more tolerant of original wording and explanation - a point taken up in the "Implications" section.

Table 2

Mean Number of Words Across Three Conditions

Group	Count	Mean (rounded)	SD (rounded)	F-ratio	F-prob
Grp 1	28	309	54	6.676	.002
Grp 2	26	326	86		
Grp 3	26	384	91		
Total	80	339	84		

The mean number of words was calculated and compared across the three groups using Anova (Table 2), and here a highly significant difference (p=.002) was found. *Post hoc* analysis (Scheffé test) shows the significant differences are between groups 1 and 3 and groups 2 and 3.

Table 3a

Mean Total T-units Across Three Conditions

Group	Count	Mean	SD	F-ratio	F-prob
Grp 1	28	19.964	4.023	2.474	.091
Grp 2	26	20.539	5.338		
Grp 3	26	23.115	6.861		
Total	80	21.175	5.598		

Scheffé test with significance level <.05= No two groups are significantly different

Table 3b

Mean Correct T-units Across Three Conditions

Group	Count	Mean	SD	F-ratio	F-prob
Grp 1	28	8.964	3.271	1.268	.287
Grp 2	26	8.000	4.775		
Grp 3	26	10.231	6.689		
Total	80	9.063	5.085		

Scheffé test with significance level <.05= No two groups are significantly different

Table 3c

Mean Incorrect T-units Across Three Conditions

Group	Count	Mean	SD	F-ratio	F-prob
Grp 1	28	11.000	4.009	1.121	.331
Grp 2	26	12.539	4.769		
Grp 3	26	12.885	5.942		
Total	80	12.113	4.956		

Scheffé test with significance level <.05= No two groups are significantly different

The mean number of t-units per script (Table 3a) did not differ significantly across groups. This suggests that the students in group 3 (control) were using longer t-units: whether this involves additional complexity and if so, how far such complexity may reside in either the clause or the nominal group (cf. Halliday 1985: xxiv) remains to be verified. Separate analyses of correct t-units (Table 3b) and of incorrect t-units (Table 3c) also indicated no significant differences at the <.05 level.

Table 4

Mean Exact Copies Across Three Conditions

Group	Count	Mean	SD	F-ratio	F-prob
Grp 1	28	1.036	1.138	3.091	.051
Grp 2	26	0.962	1.562		
Grp 3	26	2.269	3.207		
Total	80	1.413	2.197		

Scheffé test with significance level <.05= No two groups are significantly different

All t-units classified as exact copies (EC) were by definition correct. The average number of ECs was noticeably higher for the students in the control group than for the other two groups. The observed differences are very close to significance at the <.05 level (p = .0511), and the likelihood of non-random difference is sufficient to merit further study. It should be noted that students in the control group had longer to familiarise themselves with the text and extract what they considered to be the relevant parts.

Table 5a

Mean Correct Near Copies Across Three Conditions.

Group	Count	Mean	SD	F-ratio	F-prob
Grp 1	28	3.179	2.465	5.163	.008
Grp 2	26	2.769	2.372		
Grp 3	26	5.423	4.438		
Total	80	3.775	3.385		

Scheffé test with significance level <.05= Grps. 1 $ 3*; Grps. 2 & 3*

Table 5b

Mean Incorrect Near Copies Across Three Conditions.

Group	Count	Mean	SD	F-ratio	F-prob
Grp 1	28	2.214	1.912	15.523	.0000
Grp 2	26	3.808	3.418		
Grp 3	26	6.769	3.570		
Total	80	4.213	3.550		

Scheffé test with significance level <.05= Grps. 1 $ 3*; Grps. 2 & 3*

Highly significant differences were found in the mean number of near copies, both for correct near copies (NC: p = .008) and incorrect near copies (NCX: p = .0000), as shown in Tables 5a and 5b.

Taken together, the results for exact copies and near copies indicate that Group 3 students appear to have relied much more heavily on the original text in writing their summaries, and the additional time these students had for writing does not appear to have added to their accuracy. Indeed, the extent to which the different groups relied on the original text was the most noticeable difference in the students' scripts, a point taken up below.

Table 6a

Mean Correct Paraphrases Across Three Conditions

Group	Count	Mean	SD	F-ratio	F-prob
Grp 1	28	.929	1.086	.000	1
Grp 2	26	.923	1.093		
Grp 3	26	.923	1.294		
Total	80	.925	1.145		

Scheffé test with significance level <.05= No two groups are significantly different

Table 6b

Mean Incorrect Paraphrases Across Three Conditions.

Group	Count	Mean	SD	F-ratio	F-prob
Grp 1	28	1.429	1.230	3.744	.038
Grp 2	26	2.615	1.651		
Grp 3	26	2.114	1.883		
Total	80	2.038	1.657		

Scheffé test with significance level <.05= Groups 1 & 2*

As seen in Table 6, none of the groups appear to have relied heavily on paraphrases. The only significant difference across groups that was discerned at the .05 level was between groups 1 and 2 on incorrect paraphrases. The difference in distribution of text-independent elements is, rather, a consequence of the results for the category of "own wording" (Tables 7a and 7b). The mean numbers of "own wording" t-units in both instances are somewhat higher for group 1 than for group 2 and considerably higher than for group 3. The difference between groups 1 and 3 is significant at the <.05 level in both cases (p = .016 and p = .014 respectively).

Table 7a

Mean Correct Own Words Across Three Conditions

Group	Count	Mean	SD	F-ratio	F-prob
Grp 1	28	3.821	2.907	4.390	.016
Grp 2	26	3.270	3.617		
Grp 3	26	1.615	1.551		
Total	80	2.925	2.946		

Scheffé test with significance level <.05= Groups 1 & 3*

Table 7b

Mean Incorrect Own Words Across Three Conditions

Group	Count	Mean	SD	F-ratio	F-prob
Grp 1	28	1.429	7.357	4.511	.014
Grp 2	26	2.615	6.192		
Grp 3	26	2.114	4.039		
Total	80	2.038	5.900		

Scheffé test with significance level <.05= Groups 1 & 3*

From these results and from the preliminary study of individual scripts, members of group 1 (oral) after discussing the text appear to have been better able to reformulate the ideas put forward in the text into their own words. Not all their attempts at reformulation were successful. A number of the incorrect own word (OW) t-units were incorrect not simply at the grammatical level but at the conceptual level; however, this was true across groups, not only for group 1. For example:

- This gene is called "g" which is IQ tests are supposed to measure. (Group 1: script 7:5)

- ... one of the genes that plays a part in determining intelligence has tracked down. It is believed that people with such genes own more cognitive ability, measured by IQ tests. (Group 2: script 32: 7&8)

Table 8

Text Dependence/Independence Across Three Conditions (Percentages)

Group	Text Dependence (%)			Text Independence (%)		
	exact copy	near copy	total	paraphrase	own words	total
Grp 1	5.19	27.01	32.20	11.91	55.99	67.80
Grp 2	4.68	32.02	36.70	17.23	46.07	63.30
Grp 3	9.80	52.66	62.46	13.12	24.44	37.56

Table 8 summarises results according to degrees of text-dependence. For this purpose, we take exact copies and near copies to be "text-dependent" and paraphrases and "own wording" instances to be "text-independent". Values are also shown separately for our four categories. From these figures, it seems that the absence of an intervening task, associated also with more time for writing, tends to encourage outcomes that are more text-dependent.

Table 9

Mean Ideas Units (Targeted Points) Across Three Conditions

Group	Count	Mean	SD	F-ratio	F-prob
Grp 1	28	12.393	3.985	12.832	.0000
Grp 2	26	16.115	4.894		
Grp 3	26	18.039	3.550		
Total	80	15.438	4.760		

Scheffé test with significance level <.05= Groups 1 & 2*; 1 & 3*

Results for the incorporation of idea units in summaries (Table 9) show that the number of targeted points occurring in the group 1 scripts (oral task) was found to be significantly lower than in the other two groups. Although it is difficult to determine how many points needed to be targeted for an adequate summary, the findings once more suggest that having discussed the text, the students in group 1 felt somehow less obliged to adhere as closely to the original text as the students in the other two groups.

Appendix 3 gives the list of 44 targeted points. Some points such as 34 and 41 appear to have been targeted by all groups of students. Others, however, were more frequently targeted by one group than another. For example, point 35, referring to the state of mind of the children with both stimulating and supportive parents, was targeted more frequently by group 1 (oral mediating task), the frequency of occurrence being O=13, R=8, C=5. Point 44, on the other hand, concluding that research into better parenting and educational techniques will have more lasting significance, was targeted more frequently by the control group than by either of the other groups, the frequencies being O=6, R=11, C=16. Whether these differences are significant will be investigated in a later phase of the research. Further analysis of the oral discussions and the answers to the reading comprehension questions may throw some light on the points that were selected for inclusion in the summaries. At present we can only speculate that students' comments and responses did have an effect on the summaries they subsequently wrote. From overall figures, it appears likely that points covered by individuals in group 2 in the reading comprehension questions will sometimes prove to be linked to those raised in their summaries. For example, point 10 which was more frequently targeted by group 2 (the respective frequencies being O=2, R=15, C=6) was the same point raised in question 2(b) which read:

> What new kind of research technique has Plomin's team used?

to which the answer was "new gene mapping techniques".

Implications

It is worth reviewing how text-dependence or independence may be evaluated in the context of summary writing. Taylor (1984:696) reports that one of the factors that appears to distinguish professional summary writers from amateurs is the ability of the former "to detach themselves and to comment about the article." This implies that, even in summarisation, text-independence is potentially an asset. Yet teachers often appear to expect summaries to match the original text closely. In this study, for example, there appears to have been a tendency for markers to award higher marks to group 3 (control) students whose summaries were made up, on average, of over 60% more text-dependent t-units (in our terms, either exact or near copies). Subsequent discussion with the markers, and with other teachers who had carried out the activity in their classes, also revealed that these teachers believe that a good summary is one that is based closely on the original. For example, markers reportedly marked down the script (script 21, used above to exemplify "own wording") in which the student tried to compare cultivating genius with cultivating a healthy crop, on the grounds that it did not adequately summarise the passage. Another perspective on this comparison, though, could suggest that it shows evidence of reflective interaction with some of the content of the source text.

A number of teachers queried the value of the evaluative element of the task, with some suggesting that this was not a valid requirement of summary writing. It is tempting to argue that the word "summary" evokes such well established formal schemata in teachers that text-dependent answers converging on a preconceived "correct" rendering of the original content are inevitably favoured. Yet if one considers summary writing in terms of the wider academic context, then evaluating arguments becomes an integral part of summarising: academic essay writing entails the extraction and summary of selected materials from source texts as well as the evaluation of main points put forward in such source material. A possible counter to this view, in the case of work in the language class, is that genuine academic writing requires a stronger knowledge base than students in our study would have possessed in relation to the topic. While there is some truth in this reaction, we would argue that students must frequently operate from a limited knowledge base as they pick up clues or discussion material from texts they read.

As noted earlier, written products offer only indirect evidence of how learners have processed a text, and a written summary is only one form of evidence. Our paper has focused on the effects of different mediating tasks upon text processing as evidenced in final task performance, but we must appreciate that the relationship between mediating and final task will in some measure vary according to the choice of final task itself. Had our study specified the final outcome as a personal response to ideas in the source text, rather than as a summary, it is likely that teacher expectations of student writing would have differed considerably. It is also possible (though less evident) that marker reactions to source text conformity or departures from source would have privileged more text-independent outcomes than was the case in this study.

Conclusion

The initial results we have reported here already offer important practical implications for the relevance and efficiency of different treatment conditions and procedures. It is apparent that the way tasks are assigned does have an effect on students' final summary writing outcomes. If as teachers we want to encourage students to interact with the text and become less text-dependent in their summary writing, then we should encourage them to discuss the text prior to writing. However, if we want them to adhere closely to the original, then what students appear to need is time to approach the text and summary task in the way they find most suitable for the purpose.

In order to extend and strengthen the findings of the present experiment, additional work will be undertaken to compare individuals' performance on mediating tasks (contributions to oral discussion or responses to the reading exercise) with final summary outcomes. Evidence suggesting miscomprehension will also be investigated under each condition. Further studies would need to look at other students, texts, reading exercises, modes of oral discussion and final task specifications. Particular attention might profitably be given, in our view, to the effects of teacher-led or more fully prompted discussions on student participation and on subsequent summary outcomes.

Acknowledgements

This research was supported by research grant number 335/091/0001 (C.R.C.G.). We would also like to thank the students and teachers in the English Centre who participated in the study and extend our appreciation to David Churchill, Linda Cooley, Robin Corcos, Annie Mueller, Anne Storey and Denis Williamson for marking the students' scripts.

Notes

1. The limitation to written text reflects the immediate concerns of this paper. It is likely that many of the comments will also hold for listening comprehension activities.

2. On schemata, see (e.g.) several chapters in Carrell, Devine and Eskey (1988); another locally accessible source is Littlewood (1989).

3. Each treatment condition comprised two classes of students who had initially been randomly assigned to each class.

4. These judgements about "correctness" admittedly beg theoretical questions about the autonomous or derivative status of learners' own grammars and about standard English itself. This does not overly concern us since (a) judgements about correctness are routinely made by and expected of language teachers, (b) we are not claiming absolute or unique rightness for the levels of grammaticality or acceptability that informed our analyses. We believe that most of our judgements would be uncontentious among teachers (once the principle of making any such judgements is accepted), and affirm that our borderline cases were consistently resolved through the procedures we followed.

References

Alderson, J.C. & Urquhart, A.H. (eds.) (1984). *Reading in a foreign language.* London: Longman.

Allison, D.M. (1989). Helping teachers on INSET courses to read professional articles: when does intervention become interference? In Bickley, V. (ed.), (pp. 491-530).

Bickley, V. (ed.) (1989). *Language teaching styles within and across cultures.* Hong Kong: Government Printer.

Campbell, C. (1990). Writing with others' words: using background reading text in academic compositions. In B. Kroll (ed.), *Second Language Writing* (pp. 211-230). Cambridge: Cambridge University Press.

Carrell, P., Devine, J. & Eskey, D. (eds.). (1988). *Interactive approaches to second language reading.* Cambridge: Cambridge University Press.

Carrell, P. & Eisterhold, J. (1988). Schema theory and ESL reading pedagogy. In Carrell, Devine & Eskey (eds.), (pp. 77-92).

489

Davies, A. & Irvine, A. (1993). Comparing test difficulty and text readability in the evaluation of an extensive reading programme. Paper presented at the 15th LTRC, Arnhem, Holland. August 1993.

Edge, J. (1985). Do TEFL articles solve problems? *ELTJ*, 39(3):153-157.

Gunning, R. (1952). *The technique of clear writing*. New York. McGraw-Hill.

Halliday, M.A.K. (1985). *An introduction to functional grammar*. London: Edward Arnold.

Hunt, K.W. (1965). *Grammatical structures written at three grade levels*. Champaign, IL: National Council of Teachers of English.

Jenson, G.H. & DiTiberio, J.K. (1984). Personality and individual writing processes. *College Composition and Communication*, 35(3):285-300.

Littlewood, W. (1989). Constructing classroom reality. In V. Bickley (ed.), (280-292).

McCrone, J. (1993). Is there a gene for genius? In *The Independent on Sunday*, London, 2 May.

Norusis, M.J./SPSS Inc. (1992). *SPSS for Windows, Release 5.0*. Chicago, IL: SPSS Inc.

Scott, M., Carioni, L., Zanatta, M., Bayer, E. & Quintinilha, T. (1984). Using a 'standard exercise' in teaching reading comprehension. *ELTJ*, 38(2):114-119.

Stubbs, M. (1987). An educational theory of (written) language. In *British Studies in Applied Linguistics 2: Written Language*. Centre for Information on Language Teaching and Research (for British Association for Applied Linguistics).

Taylor, K. (1984). The different summary skills of inexperienced and professional writers. *Journal of Reading*, 27:691-699.

Walker, C. (1987). Individualising reading. *ELTJ*, 41(1):46-50.

Appendix 1

Text used in the study

Is there a gene for genius?

Dr Howard Gardner of Harvard University believes that geniuses are largely made. He has banned television from his home because he fears it might rot the minds of his family. He makes time everyday to listen to his seven-year-old, Benjamin, play the piano - even if it is no more than a few minutes during a transatlantic phone call while he is away at a conference.

Dr. Sandra Scarr of Virginia University, president of the Society for Research in Child Development, believes geniuses are largely born. She says parents should not worry too much about whether to take their kids to a ball game or to a museum. Talent will out.

It seems psychologists are as divided as ever over the issue of nature versus nurture. This may, however, be about to change. A conference organised earlier this year by the Ciba Foundation brought to London some of the biggest names from both sides of the debate. Startling results from unpublished work were revealed - and the beginning of a consensus could be discerned.

The most exciting results came from those working on the biology of individual differences. Dr Robert Plomin of Penn State University, working with a team from Cardiff University, hopes to announce within the next few months that he has tracked down one of the genes that plays a part in determining intelligence. An unnamed gene has been identified but the results have yet to be confirmed.

At present, it is believed that genes account for at least half of what researchers call "g" - the general cognitive ability that IQ tests are supposed to measure -while environmental influences account for the other half. But so far the evidence for a genetic component has been purely statistical, being inferred from comparisons of twins and other such hereditary studies. Plomin's method makes use of new gene mapping techniques and promises to provide direct evidence of the role that genes play.

Plomin stresses that the discovery of a first gene does not mean the riddle of intelligence has been solved. A single gene will code for only one of the many neurotransmitters and cell proteins that are the building blocks of the brain. This means that hundreds, if not thousands of genes must be involved in intelligence. The identification of even one gene does, however, have immense implications for the nature/nurture debate.

Another innovation, the computerised brain scanner, has led to a second discovery by those seeking the biological component of mental abilities. Professor Camilla Benbow of Iowa State University is head of a long-term study of the mathematically gifted. For many years she has been puzzled as to why so many of the children in her study should be boys - at the top level, boys outnumber girls by 13 to one. In a soon-to-be-published paper, Benbow reveals that the gifted boys' brains appear to process spatial information in a very different way from those of average boys and even of gifted girls.

The children in the study were scanned while being presented with a simple visual puzzle. The boys of average ability and the gifted girls showed strong activity on both sides of their brains as they thought about the puzzle. However, the gifted boys responded very differently. There was a sudden drop in activity in their left hemispheres - the side of the brain most involved in language - and an exaggerated reaction on the right, the side strongest at spatial thinking. It seems that the brains of boys with mathematical talent operate in a way that is physically distinctive.

Benbow says she was surprised that the gifted girls should lack this pattern of response. The only explanation she has is that male brains have a tendency to become more lateralised during development; when this lateralisation is taken to an extreme, unusual spatial abilities result.

Because females do not have this tendency (lateralisation is known to be hormonally governed), girls who perform well in mathematics are doing so because of a more general mental superiority. And because statistically such all-round ability is less common, this would be the reason for there being fewer mathematically gifted girls.

Benbow is quick to add, however, that cultural expectations probably exaggerate the imbalance. In China, where girls are more likely to get encouragement in mathematics, the number of gifted boys exceeds that of gifted girls by four to one rather than the 13 to one seen in the United States.

Both Plomin's and Benbow's findings would seem to give ammunition to the argument that exceptional mental abilities are largely innate. But the Ciba conference heard equally strong evidence for the role that environmental factors play in creating genius. A theme repeatedly heard from the speakers was that special children invariably have special parents.

It is a popular myth that great prodigies - the Einsteins, Picassos and Mozarts of this world - spring up out of nowhere as if touched by a divine finger. The archetype is Carl Friedrich Gauss, born into a supposedly illiterate family of labourers, who grew up to become the father of modern mathematics.

Professor William Fowler of the Massachusetts Centre for Early Learning has attacked this myth, saying that when he looked into Gauss's childhood, he found that Gauss's mother had been teaching him numerals at the age of two. His father had been a foreman, not a labourer, and played calculation games with him. Furthermore, Gauss had an educated uncle who taught him sophisticated maths at an early age.

It is the same story with other prodigies. Einstein's father was an electrical engineer who fascinatedhis son with practical demonstrations of physics. Picasso's father was an art teacher who had young Pablo copying still lifes at the age of eight. Mozart's father was a court composer who was teaching his son to sing and play almost before he could walk. "In every case, when you look into the backgrounds of great people, there is this pattern of very early stimulation by a parent or mentor figure," Fowler says.

But what sort of parental stimulation should it be? The conference heard plenty of evidence that, too often, parental pressure and attempts at "hot-housing" children result in burn-out rather than giftedness. Professor Mihaly Csiko of the University of Chicago reported on a study which identified two kinds of parental style - the supportive and the stimulating.

Supportive parents were those who would go out of their way to help their children follow their pet interests and praised whatever level of achievement resulted. Generally, such parents created a harmonious home governed by clear rules. Stimulating parents were more actively involved in what their children did, steering them towards certain fields and pushing them to work hard, often acting as a tutor.

Csiko's study followed four groups of children: one with supportive parents, one with stimulating parents, one whose parents combined both qualities and a final group who offered neither. The children were given electronic pagers; when these buzzed at random intervals during the day, they had to make a note of what they were doing and assess how happy and alert they felt.

The not too surprising result was that the children whose parents were simply supportive were happier than average but were not particularly intense in their concentration when studying or working on an interest. The children who fared best were those whose parents were both supportive and stimulating. These children showed a reasonable level of happiness and a very high level of alertness during periods of study.

Children whose parents were stimulating without being supportive were candidates for burn-out. These children did work long hours, but their alertness and happiness during study time was far below that of children in more balanced family environments.

Another crucial factor stressed at the Ciba conference is the need for parents to have proper conversations with their children. Through having the chance to talk with adults, children pick up not only language skills but also adult habits and styles of thought. One reason why prodigies such as Picasso and Einstein had a head start in life was that they had parents who demonstrated how to think about subjects like art or physics at a very early age.

Professor Fowler said a survey in Holland showed that a typical father spent just 11 seconds a day in conversation with his children. A more recent study in America produced a somewhat better result, but the fathers in question were still talking to their children for less than a minute a day.

It is not just the time spent that counts, Fowler says, but also the way in which a parent talks. A parent who brushes off a child's questions or gives dull answers will be imparting a negative, narrow-minded style of thinking. On the other hand, parents happy to take a child step by step through an argument, encouraging it to explore ideas, will foster an open and creative thinking style.

Fowler is attempting to show this experimentally with a study in which groups of parents are taught how to have constructive conversations with their toddlers. Fowler says these children have shot ahead of their peer group in language ability, intellectual ability and even social leadership skills. While the study is not yet complete, the children appear to have been given a lasting advantage.

So what is the outlook for parents who do everything right, those who manage to be both supportive and stimulating, who are good at demonstrating thinking skills to their children and successful at fostering a self-motivated approach to learning? Would such parents be guaranteed to have a gifted child?

There was general agreement at the conference that there is no denying that genuine biological differences exist between individuals; geniuses need to be lucky in both their genetic make-up and their parents. The most significant implication would seem to be that while most people are in a position to fulfil their biological potential -that is, barring serious illnesses or dietary deficiencies, they can be certain their genetic capacities will be fully developed - there can be no such certainty that they will grow up in the environment necessary for that development.

So although knowing more about the biology of genius is all very interesting, it is research into better parenting and educational techniques that will have lasting significance.

By John McCrone
The Independent on Sunday, 2/5/93
(slightly adapted)

Reprinted with permission

Appendix 2

Reading exercise used in the study

THE ENGLISH CENTRE

Student's Name: ________________________________

Number: ________________

Reading exercise: Is there a gene for genius?

Introduction

The aims of this exercise are to help you explore the text, check your understanding, look at how some of the ideas are connected, and ask some critical questions.

Different students have different needs and reactions. Please be patient if you personally find some items very easy or too difficult. We will ask for your comments later.

Your teacher will tell you how much time you have. Don't spend too long on any one item! Write your answers on the exercise sheet.

Exercise*

1. Vocabulary and ideas

a. What is the "nature/nurture" question? (Hint: If you are not sure what "nurture" means, then make a guess based on the text; it is obviously something contrasted with "nature"!).

b. Place the seven words or phrases in the list under one of the two columns. The first two have been placed for you.

LIST: 1. largely born 2. largely made 3. genetic component of intelligence

4. hereditary influences 5. environmental influences

6. innate abilities 7. cultural expectations.

"NATURE"	"NURTURE"
1	2

(* line numbers were provided for the students on their copy of the reading text for ease of reference)

2. Work by Plomin and his team

a. Note down words and phrases from the text that remind us that Plomin's results are not yet final.

b. What new kind of research technique has Plomin's team used?

3. Work by Benbow and her team

a. What new research technique was used?

b. Benbow was studying mathematically gifted children. What aspect of her study was unexpected? (Paragraph 7, lines 45-55)

c. Based on the work of Benbow and her team, answer the following questions about how gifted boys' brains work. (Hint: If the word "lateralisation" troubles you, remember that the adjective "lateral" has to do with "sides").

(i) What information do gifted boys' brains process differently from other people?

(ii) What is special about their brain activity?

(iii) Is this aspect of brain activity inborn or a result of environmental factors?

(iv) Is this aspect of brain activity the only reason that fewer mathematically gifted girls than boys are found in the United States?

Answer YES or NO _____

- Briefly explain your answer :

4. **Critical reading: read lines 86-116**

   a. "It is a popular myth that..." (line 93). Does the writer go on to agree or disagree with the belief that he reports here? ANSWER:______________

   b. Was Gauss's family illiterate?
   - Answer YES or NO_____________
   - What one word (in lines 93-98) explains your answer?_______________

   c. Fowler's comments provide reasons to suppose that the environmental role of parents (or other figures) in early life is important. What is a common factor (other than just "having special parents") in the examples he discusses (Gauss, Einstein, Picasso and Mozart)?

5. What does the use of the term "hot-housing" (line 119) tell us about the writer's attitude towards attempts by parents to make children learn and develop more rapidly?

6. a. According to Csiko's findings, which parental style or combination of styles is most beneficial for children? Circle your chosen answer.

   SUPPORTIVE STIMULATING BOTH

   b. What aspects of parent-child conversation are important, according to Fowler? (lines 152-181).

7. Do you think the conclusion (see final paragraph) is that of the discussions at the scientific conference, or the journalist writing the article?

Appendix 3

"Idea Units" targeted in the study

1. X believe that geniuses are largely born (that heredity matters) (X = Scarr, some scientists, etc.)

2. Y believe that geniuses are largely made (Y = Gardner, etc.)

3. (1 and 2 can be paraphrased as) The issue is (Psychologists are divided over) nature versus nurture.

4. The issue (3) was discussed at a recent conference (in London; Ciba foundation).

5. The beginnings of a consensus (reduction of difference) could be discerned.

6. Plomin has (probably) discovered a gene that plays a part in determining intelligence.

7. Plomin's results have yet to be confirmed.

8. At present, scientists etc. believe genes account for at least half of "g" (general cognitive ability; and think environmental factors/influences account for the other half/part).

9. Evidence (re 8) has so far been statistical.

10. Plomin used new gene mapping techniques.

11. Plomin's work promises to provide direct evidence of the role that genes play.

12. Many genes must be involved in intelligence.

13. Discovery of even one gene (that contributes to intelligence) has immense implications for the nature/nurture debate.

14. Benbow (et al) used the computerised brain scanner.

15. Benbow was studying the math. gifted (wanted to explain why so many math. gifted children were boys).

16. Benbow's study showed that gifted boys process spatial information differently.

17. (Math.) gifted boys' brains operate in a way that is physically distinctive.

18. (Benbow's explanation is that) male brains tend to become more lateralised during development.

19. Lateralisation is hormonally governed.

20. Extreme lateralisation (B. explains) results in unusual spatial abilities.

21. Girls who perform well in math. do so because of general mental superiority.

22. Cultural expectations can exaggerate the imbalance (bet. numbers of gifted boys & girls: + example comparing ratio of gifted boys to gifted girls in China and in America).

23. Plomin's and Benbow's findings strengthen the argument that exceptional mental abilities are largely innate.

24. The conference heard (equally strong) evidence for the role of environmental factors.

25. (One theme was that) special children invariably have special parents.

26. A popular myth is that geniuses just happen (divine finger!)

27. (Fowler maintained that) backgrounds of all great people had pattern of early stimulation by parent or mentor figure.

28. Examples included (some or all of) Gauss, Einstein, Picasso & Mozart.

29. (A relevant question is:) What sort of parental stimulation should it be?

30. Csiko identified (studied) two kinds of parental style - supportive & stimulating.

31. Supportive parents helped children follow pet interests (and praised whatever achievements resulted).

32. Stimulating parents pushed children towards preferred fields (and/or) pushed children to work hard.

33. Csiko's study compared (4) groups (of children) with different (combinations of) parental styles.

34. Children who fared best had parents who were both supportive and stimulating.

35. These (34) children were reasonably happy and very alert when studying.

36. Another crucial factor was the need for parents to have proper conversations with their children.

37. (Fowler said that not only amount of time but especially) the way in which a parent talks is important.

38. Parents who take a child step by step through an argument and encourage it to explore ideas will foster an open and creative thinking style (will encourage learning).

39. (When) parents (are taught to) have constructive conversations with their toddlers, these children do better (shoot ahead of peer group in language ability, leadership ability & social leadership skills).

40. There was general consensus that (no denying that) genuine biological differences exist.

41. Geniuses need to be lucky in both genetic make-up and parents (both genes and environmental factors are important).

42, Most people are in a position to fulfil their biological potential.

43. The most significant implication is that there is no certainty that the environment will provide necessary support as people develop.

44. Therefore (43) (it is) research into better parenting & educational techniques (that) will have (more) lasting significance.

THE ASSESSMENT OF SPOKEN LANGUAGE UNDER VARYING INTERACTIONAL CONDITIONS *

Vivien Berry

Abstract

Paired interviews and group discussions are becoming increasingly popular as methods of assessing spoken language. Yet recent research has shown that extreme extraverts and introverts differ in how well they perform on oral test interviews depending on whether personality types are homogeneously or heterogeneously paired. There is also experimental evidence that extraverts and introverts perform differently when tested in groups. This paper will report on a study in which approximately 100 undergraduate students were tested on their ability to take part in an academic seminar. Each student was rated by two experienced raters on a nine point scale. Ratings of speaking performance of both extremes on the extraversion scale (as measured by the EPQ) are compared to the degree of homogeneity of personality type present in each group. Initial results indicate that differences can be observed in the performances of extraverts and introverts under varying interactional conditions. The findings from this research clearly demonstrate the importance of deriving hypotheses from the psychological literature when investigating the effect of personality variables on performance. The paper concludes with a discussion of the feasibility of oral testing in groups and of the stability of results obtained.

Introduction

Of all the skills involved in learning a language, spoken language is the most difficult to assess. It is the most labour intensive and the most time consuming. Speaking is probably also the most difficult skill to score accurately and consequently scores obtained on oral language tests may not necessarily be reliable. Many factors can affect language test scores, among them differences in learners' cultural backgrounds (Chen and Henning 1985, Zeidner 1986, 1987), prior knowledge (Alderson and Urquhart 1985, Hansen and Jenson 1993), gender and academic status (Porter 1990, Cushing 1993, Zammit 1993), the extent of interviewer accommodation (Ross 1992) and different rater characteristics (Elder 1993, Pollitt and Murray 1993).

In the United States, for at least the past decade, the focus of nearly all research related to the assessment of spoken language has been the oral interview, in particular the ILR/ACTFL oral interview and its associated guidelines for the assessment of oral proficiency. This research has been primarily statistical in nature and the over-riding concern has been to provide evidence of the validity of the interview as an instrument to measure spoken language. However, the test format has been criticised for (among other reasons) not accurately reflecting the realistic features of natural communication (Bachman and Savignon 1986, Bachman 1988), or conversation (van Lier 1989).

Recognition of the shortcomings of some of the features of the interview-as-test has led official examinations organisations such as the University of Cambridge Local Examinations Syndicate (UCLES) and the Royal Society of Arts (RSA) - now amalgamated - and many university second language placement programmes to experiment with variations in both oral interview formats and oral test formats in general. One of the major innovations has been the introduction of paired or group interactions[1] between testees, rather than restricting language interaction to the traditional dyad of interviewer-interviewee. If care is taken in the allocation of learners to pairs or groups, this learner-centred approach to testing has the advantage of reducing, if not altogether removing, some of the tensions associated with the traditional dyad. For example, non-linguistic factors such as ethnicity, gender and social status, all of which have been mentioned by Brindley (1991) as potentially affecting judgements of proficiency, can be controlled for.

Without according it any special status in the hierarchy, Brindley (1991:156) also includes personality as one other non-linguistic factor amongst those he sees as important. Unfortunately, personality is a variable which cannot be controlled for on a simple observational basis. It is maintained in this paper that unless an appropriately validated instrument is used to assess personality, and allocation of learners to pairs or groups is made on a principled basis, taking into account the findings of theoretically sound empirical research, then it is misleading, to say the least, to suggest that personality has been controlled for.

Unfortunately, theoretically sound research findings into the effect of personality characteristics on second language task performance are very hard to find. The problem seems to be that specific hypotheses, derived from the specialist psychological literature, have seldom been formulated. The reason for this is that such hypotheses are not easily identified and they cannot be deduced from the second language literature. Major reviews of the role played by personality variables in second-language learning (Ellis 1986, Skehan 1989) have reached extremely pessimistic conclusions, particularly with regard to the implications of extraversion[2] as a variable. They point out that many studies have failed to produce any significant findings, citing, for example, Naiman et al. (1978), who failed to find a significant effect for extraversion in characterising the good language learner. It can be argued, however, that the problem lies not so much with the lack of significance of the results obtained but rather that these pessimistic conclusions have been reached through reviewing research which tested hypotheses that are neither logically derived from personality theory, nor predicted from relevant experimental evidence.

Another study, described by Brown as "... the most comprehensive study to date on extroversion." (Brown 1987:110) is that of Busch (1982), who also failed to find support for her somewhat extraordinary hypothesis (hunch?) that "extraverts are more proficient in English." (Busch 1982:109). More recently, Porter conducted research into affective reactions of learners based on a "rough categorisation of their personalities into 'more outgoing' or 'more reserved'..." (Porter 1991:97). Totally unsurprisingly, he also found that personality type did not seem to have any significant effect. It is findings from theoretically unsound research designs such as those of Busch and Porter, who have adapted psychological constructs merely to test things "which intuitively strike them as important" (Ellis 1986:120), that has led some researchers to reject personality as a significant factor in second language acquisition. However, summarising a comprehensive review of second-language personality studies, Griffiths concludes, "... the fact that researchers have not found relationships cannot be fairly used (as it has been) to dismiss personality variables from the L2 research agenda; nor can highly validated psychometric instruments be held accountable for the failure." (Griffiths 1991:68).

Personality measurement

The major personality dimensions are represented in almost all large scale studies and nearly all theoretical formulations. They are represented by continua, the extremes of which can be described through idealised types:

> Extraverts are sociable, like parties, have many friends and need excitement; they are sensation seekers and risk-takers, like practical jokes and are lively and active. Conversely introverts are quiet, prefer reading to meeting people, have few but close friends and usually avoid excitement. (Eysenck and Chan 1982:154)

A number of instruments have been developed which attempt to measure the major dimensions of personality, amongst them Cattell's 16PF (Cattell et al. 1970) and the Minnesota Multiphasic Personality Inventory (MMPI, Hathaway and McKinley, n.d.). Neither of these has been validated[3] for use in any non-western country and one of the major difficulties is that concepts like Introversion-Extraversion (common to all of them) which have "... an agreed meaning in one culture may not have the same, or indeed any, meaning in another culture." (Iwawaki et al. 1980:195).

When a test is used in a culture other than the one it was originally developed for, evidence of the test's reliability and validity in the new setting is required. Research has shown that reanalysis of culturally transposed tests is needed at the item level in order to identify items that function differentially for the two groups (see Ellis et al. 1993 for a detailed discussion of cross-cultural validation studies using IRT analysis). The importance of cross-cultural validation studies has been pointed out by S.B.G. Eysenck who argues that "... it is imperative that all items be tested for appropriateness before inclusion in any foreign scoring key", whilst warning of the dangers of "spurious results" if this is not done (Eysenck 1983:381)

The psychometric instruments used to assess degrees of extraversion in the studies reported here were the 86-item Japanese version of the Eysenck Personality

Questionnaire, (Iwawaki et al. 1980) and the 90 item Hong Kong EPQ (Eysenck and Chan 1982), both validated for use in the respective countries. What this means in practice is that both instruments had been subjected to translation into Japanese or Cantonese as appropriate, followed by back-translation to iron out obvious translation errors. Once translation errors had been identified and corrected, a content analysis was performed by means of inter-item correlations followed by principal component factor analysis with varimax rotation to simple structure and a final promax rotation to oblique simple structure using only the first four factors for rotation. Only items which loaded solely on one factor were included in the foreign scoring keys thus producing tests which possessed the property of measurement equivalence where "... individuals with equal standing on the trait measured by the test but sampled from different sub-populations have equal expected observed test scores" (Drasgow 1989:19).

It is appropriate to note that methods of validation of the EPQ have been criticised, most notably for the methods used to derive indices of factor comparison (Bijnen et al. 1986). Since the metric assumptions inherent in factor analysis may not be met in real data, the results, particularly of hierarchical factor analysis, may be prone to error. Non-metric multidimensional scaling,[4] which requires only ordinal assumptions of the data, offers a more robust model for multivariate analysis and may be more appropriate for analysis of the item structure of psychological tests like the EPQ. For example, two people might obtain exactly the same scores on the extraversion scale, but have achieved them by giving positive responses to different stimuli. In other words, "extraversion" is composed of more than one underlying dimension. The objective of multidimensional scaling is to determine the number of dimensions differentiating the stimuli (in this case the items in the EPQ). Individual stimuli are represented by points in geometric space; the more similar the stimuli, the closer the points. Smallest space analysis[5] of the item structure of the EPQ shows that the extraversion items form a "tight cluster" (Hammond 1987:545), thus providing further psychometric validation of the E-scale. A full discussion of the criticisms, defences and validation procedures of the EPQ is beyond the scope of this paper. However recent studies (Hanin et al. 1990) accept these criticisms and now state their results not in terms derived solely from factor analysis, but also from multidimensional scaling using smallest space analysis (Lingoes 1973).

It is clear that within the psychological community the EPQ has provoked both much criticism and a substantial body of supportive research. With the exception of the best-known IQ tests, it is probably one of the most extensively researched measurement instruments in existence. Even if philosophical doubts exist concerning the trait structure of the EPQ and the dimensions of personality it is measuring, the numerous, methodologically sound validation procedures, carried out in over forty countries over as many years, support the existence of a stable notion of extraversion which is relatively invariant, replicable and, more importantly, subject to falsification.

The problem of establishing the construct validity of the EPQ (does it measure what it is intended to measure) is, of course, circular in that there is no external criterion against which the test can be evaluated since the existence of such a criterion would make the test itself unnecessary! The only way that the validity of the EPQ can be established other than statistically, is by deriving hypotheses logically predicted from the theory, testing them and determining if they fit the predictions. A review of the experimental research reported in the psychological literature reveals several studies where it is not only possible to draw meaningful hypotheses, but also to relate them specifically to the methods of L2 testing currently under consideration. The remainder of this paper will present evidence from two such studies, the results of which show that significant differences can be observed in the responses of introverts and extraverts under varying interactional conditions.

Study 1. Paired interactions on an interview test

In the first study, extensively reported elsewhere (Berry 1993), the present researcher investigated the hypothesis, derived from Leith (1974) and further supported by the findings of Hall et al. (1988), that there would be significant differences in performance of both introverts and extraverts on an oral interview test, dependent on method of pairing. Specifically, it was predicted (again from Leith 1974) that introverts would perform best if interviewed in homogeneous pairs, next best if interviewed as individuals and worst if interviewed in heterogeneous pairs. Extraverts, on the other hand, would again perform best in homogeneous pairs but would do next best in heterogeneous pairs and worst as individuals (see Table 1). Unlike second language personality studies which generally obtain findings based on global correlational measures, psychological research in this area usually compares selected groups of extreme introverts and extreme extraverts (Cook 1993:91), "extreme" meaning plus or minus one standard deviation or more from the mean.

36 second year female students from a Japanese junior college (18 each of extreme E and I) took part in the study. No significant differences were found in their levels of general language proficiency as measured by an Institutional TOEFL. Students were randomly assigned to each of the three possible personality pairings and 24 interviews were conducted as follows (Table 2): 6 individual interviews of both I and E (12), 3 homogeneously paired interviews of both I and E (6) and 6 heterogeneously paired interviews (6). This allowed for a total of six sets of scores to be analyzed in each of the six possible categories (I-individual, E-individual, I+I, E+E, I+E, E+I). Interviews of the different categories were also conducted in random order.

Table 1

Achievements of Students Learning in Homogeneous or Heterogeneous Personality Pairs or as Individuals (from Leith 1974)

Personality	Methods		
	Homogeneous pairs	Heterogeneous pairs	Individuals
Introverts	32.2	27.3	30.0
Extroverts	30.6	27.7	25.4
Significance of Differences	n.s.	n.s.	p<.01

Homogeneous vs heterogeneous pairs: p<.01
Homogenous pairs vs individuals: p<.025

Table 2

Methods of Pairing for Interviews

Personality type	Number of interviews	Number of students
Individual I	6	6
Individual E	6	6
I+I	3	6
E+E	3	6
I+E	6	12 (6I+6E)
Total	24	36

The test itself consisted of a four part interview designed to approximate the level and format of the Cambridge Preliminary English Test (PET). Means of overall averages were calculated for each of the categories of interviews. Analysis of means yielded the following results:

Table 3

Comparison of Scores on Oral Interview Tests

Personality	**Methods**		
	Individuals	**Homogeneous pairs**	**Heterogeneous pairs**
Introverts	61.15	69.80	68.33
Extroverts	56.04	80.21	71.35
Significance of Differences	n.s.	p<.05	n.s.

Introverts: Homogeneous vs heterogeneous pairs: n.s.
Individuals vs both pairs: p<.05
Extraverts: All results p<.05

The results provide partial support for the original hypotheses. Extraverts performed exactly as predicted, showing dramatic increases over individual performance when interviewed in pairs and performing best of all in homogeneous pairs. Scores for introverts are highest in homogeneous pairs although these are not significantly different from those of heterogeneous pairs. However, against expectations, scores on individual interviews are significantly lower than on either of the pairings, suggesting that variables other than extraversion are having an effect. In fact, as both introverts and extraverts do least well in an individual interview, it may be that culturally stereotypic views of the interviewer-interviewee relationship are disturbed by, for example, having to interact in a role-play situation. Nevertheless, given the small sample size, the results are interesting and certainly indicate that further research in this area is necessary before testing in pairs is adopted wholesale.

Study 2: Participation in a group oral test.

This study investigated the hypothesis that individual learners would perform differently on a group oral test depending on the degree of extraversion of an individual in relation to the amount of extraversion present within the group.

The theoretical background for this study can be found in the work of Jennifer George (1990) who explored personality, affect and behaviour as group level phenomena in relation to absenteeism at work. She found considerable support for her hypothesis that characteristic levels of the personality traits PA (positive affect) and NA (negative affect), within work groups would be related to the positive and negative affective tones of the groups respectively. PA and NA, measured by using the appropriate scales of the Multidimensional Personality Questionnaire (MPQ, Tellegen 1982, cited in George 1990), have been shown to be related to the extraversion scale of other personality measures (George 1990:109). Characteristic levels of NA and PA within groups were determined by averaging group-member scores. Group affective tone was measured by aggregating the individual measures obtained on the Job Affect Scale (JAS, Brief et al. 1988, cited in George 1990).

Background and description of group oral test

Unlike the previous study which attempted to control as many variables as possible in an experimental research design, the setting for the current study was firmly grounded in the real world, with all the attendant constraints thus implied. The population sample was drawn from first year Economics students entering the University of Hong Kong. All first year Economics students are required by their department to take a twenty week English Enhancement course taught in the English Centre. Before the course starts, they are given an oral test, since until this year there has been no oral component in the Use of English Examination. The aim of the oral test is to provide an opportunity for students to interact with their peers in an authentic university setting , thus providing samples of language, the assessment of which provides meaningful information for both students and teachers.

The test format is designed to replicate, as closely as possible, the setting of a small academic seminar[6]. Students are assigned to groups, generally with five to a group (on the basis of their English Centre registration number which has been assigned alphabetically). They are given a short text to read, allowed five minutes to take notes on it, then asked to discuss it seriously on two levels: 1) in relation to the research and information given, and 2) by relating the research findings to their own experience and the situation in Hong Kong. Each group is assessed by two teachers - one who acts as tutor by starting off the discussion (subsequently tutors intervene only if all communication has broken down) and one who acts as an observer and takes no direct part in the proceedings. Both teachers individually assess each student using a nine point letter scale for each of: relevance / participation and articulation. At the end of each 'seminar' session, teachers discuss the grades given and agree on one grade for each category for each student.

After taking part in the oral seminar assessment exercise, each student was asked to complete a personality questionnaire. The instrument used to assess degrees of extraversion was the 90-item Hong Kong Chinese version of the EPQ (Eysenck Personality Questionnaire, Eysenck and Eysenck 1975) which emerged from the cross-cultural validation studies carried out by Eysenck and Chan in 1982 and which was kindly supplied for this study by Dr. Chan. The means and standard deviations obtained on the extraversion scale for the university sample were generally similar to those obtained by Eysenck and Chan 1982 although it will be noted that the mean for extraversion is slightly lower for the current sample.

Table 4

Comparison of EPQ Scores With Eysenck and Chan (1982)

Study	Sex	mean	s.d.	n
Eysenck and Chan 1982	(male)	12.17	4.43	270
	(female)	11.24	4.44	462
H.K.U. students 1993	(male)	11.25	4.23	38
	(female)	10.97	4.66	64

On the basis of their responses on the EPQ, students were classified as either extreme extravert, extreme introvert or ambivert. The number of extremes is interesting since the percentage in Japan was approximately 35% whereas in Hong Kong exactly 50% were classified as extremes. Obviously the higher the percentage of the population classed as extremes, the more important research is into how individual differences in personality affect performance.

Table 5

Distribution of Personality Types on Extraversion Scale of EPQ

	Extravert	Introvert	Ambivert	Total
n	26	25	51	102
mean E score	17.65	5.64	11.22	
s.d.	1.09	1.41	1.9	

As mentioned previously, students were assigned to their seminar groups quasi-randomly on the basis of their university registration numbers. The degree of extraversion present in each group was determined by averaging the E scores of each member of the group. One group consisted entirely of ambiverts (within one standard deviation of the mean in either direction) which left 20 groups for analysis. Categories of groups were then established as follows:

Table 6

Categorisation of Groups by Mean Level of Extraversion.

Category	Mean extraversion in group
1	≥ 13
2	12 < 13
3	11 < 12
4	10 < 11
5	9 < 10
6	< 9

Each category was then individually inspected to determine placement of individual extraverts and introverts within them. To control for possible differences in general language proficiency H.K.E.A. Use of English results were compared. Means were calculated for each category and analysis of means revealed no significant differences.

The results reported in Table 7 give some support to the original hypothesis that there would be observable differences in the performance of extraverts and introverts depending on the degree of extraversion present in the group. They do not, of course provide overwhelming evidence for the role of extraversion in group interaction.

However, there is a trend, supported by the finding of significant differences in the means between categories 1 and 6 that Introverts are affected by the degree of extraversion present in a group, whereas extraverts are not. It would seem that when placed in a group with a relatively high degree of extraversion, introverts respond positively to the group dynamics and therefore are rated more highly, at least for relevance/participation (no significant differences were observed between any groups for articulation) When placed in a group with a lower degree of introversion, individual introverts remain quiet and are therefore rated less highly.

Table 7

Mean Scores of Extreme E and Extreme I by Category (for Relevance / Participation)

Category	Mean Group Extraver-sion	Mean Oral score (I)	n (25)	Mean oral score (E)	n (26)
1	13≥	6.5	3	4.5	5
2	12 <13	5.6	4	5.6	6
3	11 <12	5.4	7	5.4	8
4	10 <11	5.7	4	5	7
5	9 <10	4.3	4		
6	<9	3.3	3		

Significance of differences: Extravert vs Introvert = n.s.
Extravert Groups 1-6 = n.s.
Introvert Groups 1 and 6 $p < 0.05$

Implications

The results of the studies discussed cannot be considered to provide conclusive evidence of bias either in favour of or against any particular personality type. Until they can be replicated on a much larger scale, they can only suggest potential problems of interpretation of scores. It is however interesting to note the apparently contradictory findings of the two studies with respect to the effects of extraversion. In the first study, extraverts did much better when placed in homogeneous pairs than in heterogeneous pairs whereas there were no significant differences for introverts. However, in the second study, the degree of extraversion present in a group produced no significant differences in the scores of extraverts whereas introverts were considerably affected. One possible explanation is that method has a tremendous effect on how extraverts and introverts perform. For example, there is a considerable body of research evidence which indicates that introverts are favoured by a well-structured, highly prompted learning situation (the PET is an extremely prescriptive, structured test) while extraverts are better off when presented with a high degree of uncertainty and ambiguity, such as the seminar situation (e.g. Shadbolt 1978, Riding and Parker 1979). It may be that the method effect is dominant and differences are only observed when either extreme is placed in their least favoured situation.

Given the direction towards pair and group testing by influential testing boards, this area of research could well prove to be of major importance in the very near future. Small-stakes tests are, of course, not important. Placing a student in the wrong level of class is instantly rectifiable. But what of the introverted students who turn up for the new H.K.E.A.Use of English oral exam. and find themselves placed in groups with several other introverts? What if those small differences in scores are norm-referenced so that one of them receives, for example, an E9 instead of a D8? The University of Hong Kong has an admissions policy which puts the cut-off entry point at Grade D8, so any student in the situation outlined above would be refused admission. That is when the stakes get very high indeed.

One final comment is perhaps appropriate. Even if personality characteristics are innate (and this is not altogether uncontentious), it may be that extraverts and introverts use different strategies to cope with the identical situations they both have to face in every day life. This is a very promising area for research since it adds a human dimension to the psychometric validity issues. There is at least a possibility that if differences in strategy use can be established, the problem of potential test bias can to a certain extent be overcome by appropriate learner training.

Notes

1. This is, of course, not an 'innovation' in Israel where the idea of group oral examinations dates back to at least 1980 (Reves 1980, Reves 1982, Shohamy, Reves and Bejarano 1986).

2. This paper will maintain the spelling of extraversion generally used in the psychological literature. When quoting other sources directly, the spelling used by each particular author will be adopted

3. It is important to distinguish between 'translation' by which items from a scale are translated into another language and 'validation' where items are subjected to statistical analysis (usually factor or smallest space analysis) before being included in a foreign version of a test.

4. There are two stages to multidimensional scaling. The first step is to determine the number of dimensions underlying whatever phenomenon is under investigation. The second step is to obtain scale values for the stimuli on a selected set of

dimensions. For a fuller description of the principles of multidimensional scaling and of the procedures involved, see Nunnally 1978, Chapter 2.

5. Described by Hammond (1987:544) as "One of the most elegant multidimensional scaling algorithms...", smallest space analysis was originally proposed by Louis Guttman (1968).

6. For a full description of the rationale and development of this test at the University of Hong Kong, see Morrison and Lee 1985.

Acknowledgements

The research reported in the second study was supported by research grant #335/091/0001 (C.R.C.G.). I would like to thank the students and teachers in Japan and Hong Kong who participated in these studies. In particular I would like to acknowledge the contribution of Dr. Roger Griffiths to the Japanese study. Special thanks are also due to Dr. Desmond Allison for his insightful and thought-provoking comments on earlier versions of this paper.

References

Alderson, J.C. & Urquhart, A. (1985). The effect of students' academic discipline on their performance on ESP reading tests. *Language Testing* 2(2):192-204.

Bachman, L.F. (1988). Problems in examining the validity of the ACTFL oral interview. ***Studies in Second Language Acquisition*** 10(2):149-164.

Bachman, L.F. & Savignon, S.J. (1986). The evaluation of communicative language proficiency: a critique of the ACTFL oral interview. ***Modern Language Journal*** 70(4):380-390.

Bachman, L.F., Lynch, B. & Mason, M. (1993). Investigating variability in tasks and rater judgements in a performance test of foreign language ability. Paper presented at the 15th Language Testing Research Colloquium, Arnhem & Cambridge. August 1993.

Berry, V. (1993). Personality characteristics as a potential source of language test bias. In A. Huhta, K. Sajavaara, & S. Takala. (eds.) *Language Testing: New Openings*, (pp. 115-124). Jyvaskyla, Finland: Institute for Educational Research.

Bijnen, E.J., Van der Net, T.Z. & Poortinga, Y.H. (1986). On cross-cultural comparative studies with the Eysenck Personality Questionnaire. *Journal of Cross-Cultural Psychology* 17(1):3-16.

Brindley, G. (1991). Defining language ability: the criteria for criteria. In S.Anivan (ed.) *Current Developments in Language Testing*. Anthology Series 25. (pp. 139-164). Singapore:SEAMEO.RELC.

Brown, H.D. (1987). *Principles of Learning and Teaching*. Englewood Cliffs, NJ.: Prentice-Hall.

Busch, D. (1982). Introversion-extraversion and the EFL proficiency of Japanese students. *Language Learning* 32(1):109-132.

Cattell, R.B., Eber, H.W. & Tatsouka, M.M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire (16PF)*. Champaign, Illinois: Institute for Personality and Ability Testing.

Chen, Z. & Henning, G. (1985). Linguistic and cultural bias in language proficiency tests. *Language Testing* 2(2):155-63.

Cook, M. (1993). *Levels of Personality*. (2nd edition). London: Cassell.

Cushing, S.T. (1993). L2 Proficiency, Academic Status, and Lecture Note Content. Paper presented at the TESOL Conference, Atlanta, GA. April 1993.

Drasgow, F. (1987). Study of the measurement bias of two standardized psychological tests. *Journal of Applied Psychology* 72(1):19-29.

Elder, C. (1993). Are raters' judgements of language teacher effectiveness wholly language-based? Paper presented at the 15th Language Testing Research Colloquium, Arnhem & Cambridge. August 1993.

Ellis, B.B., Becker, P. & Kimmel, H.D. (1993). An item response theory evaluation of an English version of the Trier Personality Inventory (TPI). *Journal of Cross-cultural psychology* 24(2):133-148.

Ellis, R. (1986). *Understanding Second Language Acquisition*. Oxford: Oxford University Press.

Eysenck, S.B.G. (1983). One approach to cross-cultural studies of personality. *Australian Journal of Psychology* 35(3):381-91.

Eysenck, S.B.G. & Chan, J. (1982). A comparative study of personality in adults and children: Hong Kong vs England. *Personality and Individual Differences* 3(3):153-160.

Eysenck, H.J. & Eysenck, S.B.G. (1975). *Manual of the Eysenck Personality Questionnaire*. London: Hodder and Stoughton.

George, J.M. (1990). Personality, affect and behaviour in groups. *Journal of Applied Psychology* 75(2):107-116.

Griffiths, R.T. (1991). Language Learning and the Japanese Personality: Hypotheses from the Psychological Literature. In T. Hayes & K. Yoshioka (eds.), *Proceedings of the Conference on Second Language Research in Japan*. (pp. 61-121). Tokyo:The International University of Japan.

Guttman, L. (1968). A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrica* 33:469-506.

Hall, R.H., Rocklin, T.R., Dansereau, D.F., Skaggs, L.P., O'Donnell, A.M., Lambiotte, J.G. & Young, M.D. (1988). The role of individual differences in the cooperative learning of technical material. *Journal of Educational Psychology* 80(2):172-178.

Hammond, S.M. (1987). The item structure of the Eysenck Personality Questionnaire across method and culture. *Personality and Individual Differences* 8(4):541-549.

Hanin, Y., Eysenck, S.B.G., Eysenck, H.J. & Barrett, P. (1991). A cross-cultural study of personality: Russia and England. *Personality and Individual Differences* 12(3):265-71.

Hansen, C. & Jenson, C. (1993). The effect of prior knowledge on EAP listening test performance.. Paper presented at the 15th Language Testing Research Colloquium, Arnhem and Cambridge. August 1993.

Hathaway, S.R. & McKinley, C. (n.d.). *The Minnesota Multiphasic Personality Inventory (MMPI): New Group Form (Form R)*. Minnesota: Minnesota University Press.

Iwawaki, S., Eysenck, S.B.G. & Eysenck, H.J. (1980). Japanese and English personality structure: a cross-cultural study. *Psychologia* 23(4):195-205.

Leith, G.O.(1974). Individual differences in learning: Interactions of personality and teaching methods. Personality and Academic Progress proceedings.(pp. 14-25). London:Association of Educational Psychologists.

Lingoes, J.C. (1973). *The Guttman-Lingoes nonmetric program series*. Ann Arbor: Mathesis.

Morrison, D.M. & Lee, N. (1985). Simulating an academic tutorial: a test validation study. In Y.P.Lee, A.C.Y.Y. Fok, R.Lord & G.Low (eds.) *New Directions in Language Testing*, (pp. 85-92). Oxford: Pergamon Press.

Naiman, N., Frohlich, M., Stern, H.H. & Todesco, A. (1978). *The Good Language Learner. Research in Education Series #7*. Toronto: Ontario Institute for Studies in Education.

Nunnally, J.C. (1978). *Psychometric theory*. (2nd edition) New York: McGraw Hill.

Pollitt, A. & Murray, N.L. (1993). What raters really pay attention to. Paper presented at the 15th Language Testing Research Colloquium, Arnhem and Cambridge. August 1993.

Porter, D. (1991). Affective factors in the assessment of oral interaction: gender and status. In S.Anivan (ed) *Current Developments in Language Testing*. Anthology Series 25. (pp. 92-102). Singapore: SEAMEO.RELC.

Reves, T. 1980. The group-oral Bagruth examination: an experiment. *English Teachers Journal* 24. (quoted in Reves 1982).

Reves, T. 1982. The group-oral examination: a field experiment. *World Language English* 1(4):259-262.

Riding, R.J. & Parker, J.E. (1979). The effect of extraversion, detail importance and interference on the recall of prose by eleven-year old children. *Educational studies*, 5(1):15-22.

Ross, S. (1992). Accomodative questions in oral proficiency interviews. *Language Testing* 9(2):173-186.

Shadbolt, D.R. (1978). Interactive relationships between measured personality and teaching strategy variables. *British Journal of Educational Psychology*, 48(2):227-231.

Shohamy, E., Reves, T. & Bejarano, Y. (1986). Introducing a new comprehensive test of oral proficiency. *ELT Journal* 40(3):212-220.

Skehan, P. (1989). *Individual Differences in Second-Language Learning*. London: Edward Arnold.

Stansfield, C. & Hansen, J. (1983). Field dependence-independence as a variable in second language cloze test performance. *TESOL Quarterly* 17(1):29-38.

van Lier, L. (1989). Reeling, writhing, fainting and stretch in coils: oral proficiency interviews as conversation. *TESOL Quarterly* 23(3):489-508.

Zammit, S.A. (1993). Motivation, test results, gender differences and foreign languages: how do they connect? Paper presented at the 15th Language Testing Research Colloquium, Arnhem and Cambridge. August 1993.

Zeidner, M. (1986). Are English language aptitude tests biased towards culturally different minority groups? Some Israeli findings. *Language Testing* 3(1):80-95.

Zeidner, M. (1987). A comparison of ethnic, sex and age biases in the predictive validity of English language aptitude tests: some Israeli data. *Language Testing* 4(1):55-71.

THE GRAMMATICAL AWARENESS AND KNOWLEDGE OF HONG KONG TEACHERS OF ENGLISH

Stephen Andrews

Introduction

This paper reports findings from a preliminary investigation into the grammatical awareness and grammatical knowledge of Hong Kong teachers of English and into their attitudes towards grammar.

Selection of this particular focus was prompted by a number of factors:

1. Whatever one's view of the value of explicit grammar teaching, the learning or acquisition of grammar remains central to the study of language;
2. In Hong Kong schools, despite the official adoption of a communicative approach to language teaching, explicit teaching of grammar has continued unabated.

Research into the level of grammatical awareness and knowledge that Hong Kong teachers of English bring to the teaching task seems especially timely in view of the current concern about the proportion of teachers employed as English teachers in local secondary schools who do not possess a relevant degree. A 1989 Education Department report estimated that 46% of teachers of English are not subject-trained (education Department 1989:60). It is, of course, a matter of debate as to what constitutes appropriate subject-training for such teachers. It could be argued, for instance, that a degree in English which has concentrated mainly on the study of literature is not necessarily an adequate basis for a career as a teacher of English language. Nevertheless, the underlying worries about the possible consequences of a lack of subject-knowledge seem perfectly understandable and worthy of further investigation.

Background

In the 1980s, in the UK in particular, considerable attention began to be given to 'Language Awareness.' This was partly in response to the ideas proposed by Hawkins and others in the early 80s that 'Language Awareness' should form part of the school curriculum: that explicit knowledge of forms of language, the structure and development of language, language in use, and L1 and L2 acquisition might help, among other things, to provide a bridge between the teaching of the mother-tongue and of foreign languages (Hawkins 1981, 1984, Donmall 1985). Then, in the mid- to late 80s, a succession of DES reports on the teaching of English as a mother-tongue, beginning with *English 5 to 16: Curriculum Matters*

1 (DES 1984) and culminating with the so-called Kingman report (DES 1988), argued for pupils to be taught more directly about the forms and structures of the English language.

These reports served to fuel the growing interest in language awareness. A 1989 British Association for Applied Linguistics Seminar took language awareness as its theme. Then, a year after the publication of selected papers from that BAAL Seminar (James and Garrett 1991a), the journal *Language Awareness* was launched, with the declared aim to encourage and disseminate work which explores:

1. The role of explicit knowledge about language in the process of language learning;

2. The role that such explicit knowledge about language plays in language teaching and how such knowledge can best be mediated by teachers;

3. The role of explicit knowledge about language in language use.

In the last few years, in part resulting from the debate surrounding the Kingman report and the associated LINC (Language in the National Curriculum) teacher-training project, there has been increased interest in and research into the language awareness of practising teachers, particularly of English as a mother-tongue and of foreign languages (see, for example, Mitchell and Hooper 1991). In relation to teachers of English as a foreign language, by contrast, there has been relatively little published research (though see, for example, McNeill's work on vocabulary and also Andrews, in press).

The present study focuses specifically on teachers' awareness and knowledge of grammar, but what exactly is understood by these terms in the context of this discussion? In Andrews, in press, trainers of English native-speaker teachers of EFL were asked to characterise the grammatical knowledge and awareness of teachers. Table 1 indicates the range of aspects mentioned.

The variety of responses of the trainers reflects the multifaceted nature of language awareness (see, for example, Stainton 1992) and illustrates something of the complexity surrounding any attempt to define what it is teachers of EFL should know/understand/be aware of/be able to do in relation to grammar. Clearly some of the points raised relate more to grammatical knowledge, others to grammatical awareness, while some seem to involve aspects of technique as well.

James and Garrett discuss five domains of language awareness: the affective, social, 'power', cognitive and performance domains (James and Garrett 1991b). The results discussed in the present paper focus principally on the cognitive and performance domains - grammatical knowledge in the sense of being able to understand and apply grammatical terms correctly - although reference will also be made to the affective domain and responses throwing light on:

1. Teachers' views of the importance for teachers and learners of knowing grammatical rules and terminology;
2. Teachers' confidence in their own knowledge of grammar/about grammar;
3. Teachers' attitudes towards grammar as learners of a language.

Table 1

Components of EFL Teachers' Grammatical Knowledge Awareness: the Views of Trainers

What Characterizes Grammatical Knowledge/Awareness?

1. Knowledge of grammatical terminology
2. Understanding of the concepts associated with terms
3. Awareness of meaning/language in communication
4. Ability to reflect on language and analyze language forms
5. Ability to select/grade language and break down grammar points for teaching purposes
6. Ability to analyze grammar from learners' perspective
7. Ability to anticipate learners' grammatical difficulties
8. Ability to deal confidently with spontaneous grammar questions
9. Ability to think on one's feet in dealing with grammar problems
10. Ability to explain grammar to students without complex metalanguage
11. Awareness of 'correctness' and ability to justify an opinion about what is acceptable usage and what is not
12. Sensitivity to language/awareness of how language works

Methodology

The research strategy adopted in this instance was that of the questionnaire. It was felt that this would enable a wide range of information to be collected from a large number of respondents, with the possibility that their responses would bring to light a variety of interesting possibilities for further, more in-depth study.

The questionnaire was administered to 141 teachers and prospective teachers of English, the majority of whom were enrolled on courses of initial teacher-training. Of the 141 respondents, 101 were native-speakers of Cantonese, 29 were native-speakers of English, 9 were native-speakers of other languages, while two failed to specify their native language. The native-speakers of Cantonese were in the main practising teachers in Hong Kong secondary schools, but without much experience, while the native-speakers of English (most of whom were based in the UK) generally had no teaching experience.

The response rate from the Cantonese native-speaker teachers was 100%, because the questionnaire was administered to them during one of their classes at the University of Hong Kong. The English native-speakers, by contrast, were a self-selecting group. Therefore, they may well not be representative of English native-speaker teachers of EFL as a whole. This hypothesis would appear to be borne out by the comment of one such respondent, written at the bottom of the questionnaire: "Many people didn't do this test as they felt it would be a test of their grammar, which they didn't feel too confident of, so maybe it's not a very representative sample." The suggestion that many English native-speaker trainee teachers of EFL lack confidence in relation to grammar accords very closely with the experience of their trainers, whose responses indicated that "...a large number of native-speakers begin their careers as EFL teachers with a marked lack of confidence in their own knowledge/awareness of grammar, experiencing feelings variously described as insecurity, inadequacy, fear and panic" (Andrews, in press).

Before discussing the results of the questionnaire, the design should be described. The questionnaire was made up of four sections:

1. Section A contained a number of personal detail questions eliciting data about respondents' language background, educational background and teaching experience.

2. Section B consisted of five open-ended questions designed to explore respondents' awareness and understanding of what grammar is (e.g. What are *rules* of grammar? What does it mean for a *learner* of English to 'know' a grammatical rule? What does it mean for a *teacher* of English to 'know' a grammatical rule?). The results from this section have yet to be analyzed.

3. In section C respondents were asked to consider a number of questions or statements and to record their replies using a five-point scale. The first set of questions elicited views on the importance of grammatical knowledge for teachers and for learners. A response of 5 represented "very important", while a 1 indicated a "not at all important" rating. The second group of questions elicited respondents' assessments of their confidence in their own grammatical knowledge (5 = very confident, 1 = not at all confident), while the third focused on their attitudes to grammar as learners of a language (5 = strongly agree, 1 = strongly disagree).

4. Section D represented an attempt to obtain information about respondents' understanding of grammatical terms and their ability to apply them correctly. The first two parts of this section were taken directly from a questionnaire developed by Bloor (see Bloor 1986a, 1986b) and known as the SPAM questionnaire (Students' Prior Awareness of Metalinguistics). Bloor's original questionnaire was administered to undergraduates in two British universities. One group of 63 undergraduates - referred to by Bloor as the

'linguists' - were just entering the first year of modern languages or linguistics programmes, while the other group of 175 - the 'non-linguists' - were second-year students of other subjects, but with some interest in foreign language study. In the first of the items borrowed from the SPAM questionnaire, respondents were given a sentence and were asked to locate within it one example of each of fifteen different parts of speech (e.g. countable noun, adverb, preposition), while in the second they were given four sentences and were asked to identify in each the word or phrase performing a specified grammatical function (e.g. subject, direct object). The third part of Section D was a task requiring respondents to supply appropriate metalanguage themselves rather than matching a given item to its exemplar, by providing labels for a number of different verb forms within a short piece of text. (See Appendix A for sections B, C and D).

Results

The first part of this section discusses the responses of the Cantonese native-speakers in comparison with those of the English native-speakers. In considering these results, it should be kept in mind that, as mentioned earlier, the English native-speaker group is probably not representative, a hypothesis which needs to be tested in further research.

The first set of responses (see Table 2) shows how the two groups rated the

Table 2

The Importance of Grammatical Knowledge

	English NS Mean	English NS s.d.	Cantonese NS Mean	Cantonese NS s.d.
1. Important to know grammar rules for a teacher	4.69	(0.54)	4.7	(0.50)
2. Important to know grammar rules for a learner	3.48	(0.91)	4.15	(0.89)
3. Important to know grammar terms for a teacher	4.55	(0.57)	4.27	(0.77)
4. Important to know grammar terms for a learner	2.86	(1.51)	3.01	(0.93)
5. Important to teach grammar to learners of English	3.59	(1.05)	4.18	(0.90)

importance of grammatical knowledge. Both groups considered knowledge of grammar rules and terms to be important for a teacher. Rules were felt to be more important than terms by both groups for both teachers and learners, while the English native-speakers gave greater importance to teachers' knowledge of terms than the Cantonese native-speakers did. One of the most interesting contrasts between the two groups of respondents can be seen in their assessments of the importance of grammatical knowledge for learners: both groups see the teaching and learning of grammar as important, but the Cantonese native-speaker respondents provide consistently higher ratings.

The responses set out in Table 3 show the levels of confidence expressed by the two groups. As one might expect, the responses of the English native-speakers reveal that the members of this group have greater confidence in their ability to speak and write correct English. The English native-speakers indicate that they have marginally more confidence in their speaking ability than in their writing ability, while the Cantonese native-speakers show more confidence in writing than in speaking. Otherwise, the levels of confidence are very similar, with the Cantonese native-speakers no less confident than their English native-speaker counterparts. The relatively high levels of confidence expressed by the former group are perhaps

Table 3

Levels of Teachers' Confidence

	English NS		Cantonese NS	
	Mean	s.d.	Mean	s.d.
1. Ability to **speak** correct English	4.62	(0.49)	3.76	(0.75)
2. Ability to **write** correct English	4.59	(0.50)	3.99	(0.69)
3. Knowledge of rules of English grammar	3.52	(0.57)	3.6	(0.74)
4. Knowledge of grammatical terms	3.41	(0.68)	3.37	(0.82)
5. Ability to explain a grammatical rule correctly	3.24	(0.87)	3.38	(0.87)
6. ability to think of good examples to illustrate a grammatical rule	3.48	(1.02)	3.34	(0.83)

rather less surprising than the similarly high levels expressed by the latter group, in view of the feelings of insecurity and panic referred to earlier. This would tend to reinforce the view that this particular group of English native-speaker trainee EFL teachers may be an unrepresentative sample.

The attitudes towards grammar of these two groups when learning a foreign language reveal some interesting points of comparison, as can be seen in Table 4. The contrasts contained in the first two sets of responses are possibly as one would have predicted, with the English native-speaker group paying less attention to grammar than the Cantonese native-speakers and the latter group finding it more useful to learn grammar rules by heart than the former. The fifth set of responses also shows a contrast which one might have anticipated, with the English native-speakers agreeing more strongly with the idea that it might be more useful to practise communication than to spend time on grammar. What is perhaps slightly surprising about this set of responses is the relatively high level of agreement expressed by the Cantonese native-speakers. The contrasting attitudes revealed in the third and fourth sets of responses were possibly less predictable: the English native-speakers show a stronger preference for grammatical explanations from the teacher, while the Cantonese native-speakers seem to have the greater enthusiasm for an inductive approach.

Table 4

Attitudes as Language Learners

	English NS Mean s.d.	Cantonese NS Mean s.d.
1. I pay very little attention to grammar.	3.21 (1.90)	2.22 (1.25)
2. I find it useful to learn grammar rules by heart.	3.07 (1.53)	3.63 (1.15)
3. I like my teacher to give grammatical explanations.	4.03 (1.35)	3.89 (1.08)
4. I like to look at examples and work out rules by myself.	3.34 (1.74)	3.72 (1.20)
5. I find it more useful to practise communication than to spend time on grammar.	3.79 (0.86)	3.22 (1.21)
6. I am very interested in the grammar of that language.	3.38 (1.18)	3.48 (1.15)

In those tasks which tested respondents' understanding of grammatical terms and their ability to apply them correctly, the Cantonese native-speakers performed better on two tasks out of three, those with the larger number of items and therefore likely to provide more reliable results (see Table 5).

Table 5

Knowledge of Grammar/Grammatical Terminology

	English NS Mean s.d.	Cantonese NS Mean s.d.
1. Ability to identify specific parts of speech (max. 15)	12.54 (2.30)	12.89 (3.47)
2. Ability to identify grammatical functions/relations (max. 4)	2.88 (0.93)	2.64 (0.70)
3. Ability to name verb forms (max. 30)	26.59 (3.72)	27.97 (5.29)

Again, one suspects that the difference in the performance of the two groups would have been greater if the English native-speakers had been a more representative sample, a suspicion which is lent support when one compares the performance of both groups with the performance of Bloor's respondents on the two common tasks (Table 6).

If one looks at the two left-hand columns first, perhaps the most striking feature is the generally poor performance of the English native-speaker non-linguists in Bloor's sample. On every item their results are worse than those of the English native-speakers in the present study, in many cases dramatically so. Given that a significant proportion of English native-speakers entering the TEFL profession are not subject specialists, these figures would tend to confirm that the present sample is not representative.

Across the two samples there are a number of interesting points of comparison including, for example, those where both Bloor's groups perform markedly worse than the groups in the present study (countable noun, adverb, definite and indefinite articles) and those where Bloor's non-linguists stand out (past participle, conjunction). Perhaps the most noteworthy feature of all is the generally good performance of the Cantonese native-speakers, especially since the sample includes a significant number of non-specialists.

Table 6

SPAM Questionnaire Results

(Comparison of Bloor (1986) and present study)

	Ling	N-Ling	Eng NS	Canton-ese NS
verb	0	5	0	2
noun	0	7	3	11
countable noun	38	43	7	12
passive	27	85	45	23
adjective	2	27	10	12
adverb	25	66	7	18
definite article	21	48	7	14
indefinite article	33	67	7	18
preposition	9	60	0	10
relative pronoun	17	68	10	18
auxiliary verb	44	54	38	46
past participle	21	45	14	17
conjunction	11	55	10	14
finite verb	41	56	48	42
infinitive	11	58	14	34
subject	0	4	0	1
predicate	88	92	65	73
direct object	8	42	7	3
indirect object	30	46	45	61

The second part of this section looks a little more closely at the Cantonese native-speaker respondents, i.e. the Hong Kong secondary school teachers of English. Three comparisons between different sub-groups within the sample of 101 were made to see if any interesting points of contrast emerged.

The first of these comparisons was between those respondents who had received their tertiary-level education in Hong Kong (of whom there were 55) and those who had studied at overseas universities (numbering 33). As can be seen from Table 7, those who studied in Hong Kong performed a little better than their overseas-study counterparts in all those tasks testing their understanding of and ability to apply grammatical terms correctly.

The second comparison was between those whose subject of study at tertiary level was felt to be relevant to teaching English as a foreign/second language (of whom there were 48) and those whose university studies were in an area considered not to be relevant (33 respondents). For the purposes of this comparison, those respondents who said they had studied English or linguistics were placed in the first category. In an attempt to reduce 'fuzziness' and sharpen any potential contrasts, those who specified Education as their subject of tertiary study were excluded from either group. Nevertheless, an element of fuzziness inevitably remains, since the profile of the 'specialists' excludes such information as exactly how much English they studied, and what sort of English they studied (did their course, for example, focus mainly upon the study of literature?).

Table 7

Cantonese NS Teachers: Place of Tertiary Study

	Study in Hong Kong (n = 55) Mean s.d.	Study Overseas (n = 33) Mean s.d.
1. The importance of grammar	20.07 (2.44)	20.35 (2.95)
2. Confidence re. grammar	21.51 (3.55)	21.58 (3.92)
3. Ability to identify specific parts of speech	13.57 (2.37)	12.43 (3.79)
4. Ability to identify grammatical functions/relations	2.69 (0.71)	2.48 (0.63)
5. Ability to name verb forms	28.64 (4.73)	27.32 (5.30)

Table 8 shows the results of this second comparison. Not surprisingly, those whose tertiary studies included English or linguistics had noticeably more confidence in their grammatical knowledge. Their results were also better on two of the three tasks focusing on grammatical terms, although on the first of the three (that requiring respondents to identify specific parts of speech) they performed less well than the non-specialists.

The final comparison looked at respondents' years of teaching experience. Two groups were identified from within the original sample: those with up to one year's full-time experience of teaching English (of whom there were 22) and those with at least six years of experience (24 in number). As can be seen in Table 9, those with at least six years' experience were both more confident and generally better in their understanding of and ability to apply grammatical terms correctly. Once more, however, there is an element of fuzziness surrounding these results, since those respondents with at least six years' experience had also in the majority of cases already followed a course of initial teacher-training, while the less experienced group had not.

Table 8

Cantonese NS teachers: Subject of Tertiary Study

	Studies included English/linguistics (n = 48)		Studies not relevant to TEFL (n = 33)	
	Mean	s.d.	Mean	s.d.
1. The importance of grammar	20.13	(2.78)	20.13	(2.47)
2. Confidence re. grammar	22.19	(3.89)	20.61	(3.58)
3. Ability to identify specific parts of speech	12.82	(3.71)	13.25	(2.82)
4. Ability to identify grammatical functions/relations	2.80	(0.75)	2.36	(0.49)
5. Ability to name verb forms	28.92	(3.13)	27.50	(6.58)

Table 9

Cantonese NS Teachers: Years of Teaching Experience

	Up to 1 year's experience (n=22)		Over 6 years' experience (n=24)	
	Mean	s.d.	Mean	s.d.
1. The importance of grammar	19.68	(2.98)	20.26	(2.51)
2. Confidence re. grammar	20.73	(3.79)	23.25	(3.80)
3. Ability to identify specific parts of speech	12.59	(3.30)	12.71	(3.33)
4. Ability to identify grammatical functions/relations	2.42	(0.51)	2.70	(0.76)
5. Ability to name verb forms	25.81	(5.97)	28.45	(6.19)

Conclusion

Probably the first conclusion to be drawn from a preliminary study of this nature is that, because the results are tentative, one should not in fact draw any very firm conclusions beyond the need for much more research.

The comparison between the English native-speaker and the Cantonese native-speaker teachers of EFL reveals some interesting points of contrast. Although it has been suggested that the English native-speaker sample is not representative, the Cantonese native-speaker sample, by contrast, is considerably larger and rather more representative. It would therefore not be unreasonable to take the profile of attitudes, knowledge and awareness presented by this category of respondents to be fairly typical of Hong Kong secondary school teachers of English.

The differences between sub-groups within the population of Cantonese native-speaker teachers point to certain trends with, for example, the subject of tertiary study and the number of years of teaching experience appearing to have quite a marked effect upon a teacher's level of confidence in relation to grammar. One is left, however, with the realisation that the contrasts do not fall into conveniently neat patterns, that no single variable does consistently have a significant effect upon grammatical knowledge/awareness. There is clearly a great deal of variation between individual teachers, each of whom is an amalgam of different characteristics and the product of a range of linguistic and educational experiences, any of which, singly or in combination, may have had some impact upon that individual's grammatical knowledge and awareness.

Therefore, the final conclusion to be drawn from this preliminary study is that there is a need for in-depth research which looks more closely at a much smaller number of teachers. This research should aim not only to analyze what the subjects understand by grammar, grammar rules and 'knowing' grammar rules, and to examine their level of grammatical knowledge and awareness, but also, perhaps most interestingly, to investigate how their grammatical knowledge and awareness (or lack of it) impacts upon the classroom.

References

Andrews, S.J. (In press). The grammatical knowledge/awareness of native-speaker EFL teachers - what the trainers say. In M. Bygate, A. Tonkyn & E. Williams (eds.). *Grammar and the language teacher.* Hemel Hempstead: Prentice Hall.

Bloor, T. (1986a). University students' knowledge about language - some aspects of language awareness prior to instruction in university courses. *CLIE Working Papers,* Number 8.

Bloor, T. (1986b). What do language students know about grammar? *British Journal of Language Teaching,* 24:3.

Department of Education and Science (DES) (1984). *English 5 to 16:curriculum matters 1*. London: HMSO.

Department of Education and Science (DES) (1988). *Report of the Committee of Inquiry into the teaching of English language*. London: HMSO.

Donmall, B.G. (ed.) (1985). *Language awareness: NCLE Reports and Papers*, 6. London: CILT.

Education Department (1989). *Report of the Working Group set up to review language improvement measures*. Hong Kong: Hong Kong Government.

Hawkins, E.W. (1981). *Modern languages in the curriculum*. Cambridge: Cambridge University Press.

Hawkins, E.W. (1984). *Awareness of language: an introduction*. Cambridge: Cambridge University Press.

James, C. & Garrett, P. (eds.) (1991a). *Language awareness in the classroom*. London: Longman.

James, C. & Garrett, P. (1991b). *The scope of language awareness*. In C. James & P. Garrett (eds).

McNeill, A. (1993). New language, content and context: some sources of confusion for ESL learners. In N. Bird, J. Harris & M. Ingham (eds.) *Language and Content* (pp. 462-470). Hong Kong: Government Printer.

Mitchell, R. & Hooper, J. (1991). *Teachers' views of language knowledge*. In C. James & P. Garrett (eds).

Stainton, C. (1992). Language awareness: genre awareness - a focused review of the literature. *Language Awareness* 1:2.

SOME CHARACTERISTICS OF NATIVE AND NON-NATIVE SPEAKER TEACHERS OF ENGLISH

Arthur McNeill

Introduction

It is often assumed that teachers who teach their own mother tongue have a number of advantages over teachers who are not native speakers of the language they teach. Native speaker intuitions about language are supposed to result in the production of correct, idiomatic utterances, as well as providing the ability to recognise acceptable and unacceptable versions of the language. Most non-native speaker teachers of English can only aspire to this. However, a possible disadvantage faced by the native speaker teacher of ESL is the linguistic distance between teacher and learner. Are native English speakers likely to be less sensitive to their learners' language needs because they have less access to their students' language and, by extension, to the way in which their students process English as a second language?

This paper reports the results of a study which compares the performances of four groups of Hong Kong teachers on a language teaching task. Two of the groups consist of native English speakers, while the others consist of native speakers of Cantonese, which is the L1 of the students referred to in the study. The difficulties involved in defining precisely what constitutes a native speaker of a language have long been of interest to applied linguists. Davies (1991) argues that the differences are far from clear-cut and that there is the possibility of mobility from non-native to native speaker. However, as far as the subjects in the present study are concerned, none of the non-native speakers of English would wish to be considered as native speakers. While all of them grew up in Hong Kong and received most of their education there through the medium of English, their dominant language is Cantonese. Their use of English tends to be restricted to school and university, with Cantonese being used almost exclusively at home and on social occasions.

Background

The study addresses the issues of teachers' language awareness and their sensitivity to students' language difficulties, in particular the ability to anticipate the problems which learners face when exposed to particular texts. Brutten (1981) argues that ESL teachers are generally accurate in their anticipation of students' vocabulary difficulties. Her conclusions are based on a study which asked a group of native speaker ESL teachers to identify the difficult vocabulary items in a reading text intended for a particular level of ESL learner. A group of ESL students of the same proficiency level was asked to identify the words which they found most difficult in the text and the patterns of teacher and student selections were compared. The words identified by the teachers were largely the same as those selected by the students. When the Brutten study was replicated in Hong Kong (McNeill 1992), using Cantonese-speaking teachers and students, it was found that while the correlation between the teachers' and students' selections was high and indeed close to the correlation established by Brutten (r=.674, p<.01, Brutten 1981; r=.664, p<.001, McNeill 1992), some important differences emerged in the patterns of choice between the two groups. Since correlation is a trend analysis, it is not possible to conclude from the high correlations alone that the patterns of choice of the two groups were significantly similar. In order to determine whether there was any significant difference between the two groups' results, Brutten then ran a t-test, the results of which were not significant. By contrast, the results of a t-test on the Hong kong samples was highly significant (t=3.75, p<.0001), which suggested that differences existed in the patterns of choice, which merited closer investigation.

An obvious difference in the two studies lies in the use of native and non-native speaker teachers. Did the ESL teachers who were native English speakers have some sort of advantage as a result of their native speaker intuitions about English? In order to explore this line of investigation further, the present study compares groups of native and non-native speaker teachers on the same language task.

A possible weakness of the two studies mentioned above is that there is no objective measurement of the students' vocabulary knowledge. The students simply underlined words in a text to indicate that they required explanation of the meaning. It is quite possible that the studies included students who over-estimated their vocabulary knowledge, e.g. by guessing wrongly the meaning of some words, as well as some students who under-estimated their vocabulary knowledge, e.g. by claiming they did not know the meaning of a word when they had inferred the meaning correctly from the context. McNeill's (1993) follow-up study introduced an objective vocabulary assessment. A 40-item vocabulary test was constructed, consisting of the 40 words which a sample of teachers identified as the most likely to present difficulty for ESL students reading the text. The vocabulary tests were administered to 200 upper-intermediate level (Form 6) secondary school students. The tests were given to each student twice: first as a list of 40 isolated words for which the student had to provide the L1 (Chinese) equivalent, then in connection with the reading text from which they were taken, allowing the students to consult the text and make use of contextual clues. The teachers' selections of words were then compared with the results of the two vocabulary tests. The extent to which teachers identified the difficult words in the second (contextualised) vocabulary test became the main focus of interest, since the ability to identify contexts which provide learners with clues to the meaning of unknown words is an essential part of the teachers 'sensitivity to learners' difficulties.

The ways in which ESL learners cope with contextual clues have been investigated by Liu and Nation (1985) and Li (1988). The characteristics of contexts which are helpful to language learners have also been examined (Perkins and Brutteh 1983, Schouten van Parreren 1985). Since the teachers who took part in the study were meant to identify only the words which the students needed to know to understand the main ideas in the text, it was not intended that the students should learn new vocabulary from inferring the meaning of new words. Indeed, researchers are cautious about recommending inferring from context as a means of learning new vocabulary (Mondria, J.-A. and Wit-de-Boer 1991).

Research Questions

The present study addresses three questions:

1. Do (native speaker) ESL teachers, as a group, predict learners' vocabulary difficulties in reading texts accurately?

2. To what extent does the ability to predict learners' vocabulary difficulties vary among individual teachers?

3. What similarities and differences can be detected in the way native and non-native speaker teachers of English predict learners' vocabulary difficulties in reading texts?

Subjects

Four groups of teachers took part, two of whom were native speakers of English and two of whom were native speakers of Cantonese. The two groups for each language consisted of an experienced ("Expert") group and an inexperienced ("Novice") group. The Expert group of Cantonese speakers consisted of teachers who had completed a degree, a postgraduate qualification in English teaching and were, at the time of the study, in the process of following a master's degree in teaching ESL. Not only were they advanced users of English, as far as their own proficiency was concerned, but they also had a good knowledge of language analysis and had taken courses in language acquisition. The Expert group of English native speakers had all completed a relevant master's degree. All of the Experts had at least three years' experience of teaching ESL. By contrast, the Novice groups had had little exposure to education theory, language teaching methodology and very limited experience of teaching ESL. The Cantonese speaking group were all English majors in the first year of studying for a bachelor's degree in education. Their theoretical knowledge of language analysis and language acquisition was basic. The English native speaker Novices were attending the Royal Society of Arts Certificate in Teaching English as a Foreign Language at the British Council in Hong Kong. This is a pre-service course intended for teachers interested in embarking on a career in TEFL. None of the English native speakers in the study had more than a very basic command of Cantonese.

The four groups of teachers can be characterised as follows:

1. Native Speaker Expert (NSE)

2. Native Speaker Novice (NSN)

3. Non-native Speaker Expert (NNSE)

4. Non-native Speaker Novice (NNSN)

The use of the terms "Expert" and "Novice" to describe the two levels of teacher expertise is based on Berliner's (1989 and forthcoming) theory of the development of expertise in teachers. In the present study only two extremes of teacher expertise are included. Berliner's research compares the performance and attitudes of teachers across various levels of expertise: "novice", "advanced beginner", "competent", "proficient" and "expert".

The students are the 200 Form 6 pupils (Arts Stream) from Hong Kong secondary schools who took part in the vocabulary tests in McNeill 1993. The text used is the same as in the previous study. The students' scores on the two 40-item vocabulary tests mentioned above are also used in the present study.

Procedure

The four Teachers' groups were asked to preview the text, *The Sword That Can Heal* (Appendix) and to imagine that they would use the text in a reading skills lesson with Form 6 (Arts Stream) students (or Hong Kong students of equivalent English proficiency). They were asked to select the twelve words from the text which, in their judgement, the students would need to have explained to them in order to gain a general understanding of the text. It was assumed that the words selected would be new to the students and that the students could not be expected to infer the meanings of the words from the context. In making their selection, the teachers had to ask themselves the following questions:

1. Which words were unfamiliar to the students?
2. Which were essential for a grasp of the general sense of the text?
3. How many of the new words could the students deal with on their own, using contextual or word-internal clues?

Data Analysis

The data analysis was conducted as follows:

1. The patterns of four Teacher group selections were compared with the results of two Students' vocabulary tests (correlation).
2. Individual differences within the four Teachers' groups were analyzed (scatterplot, phi coefficient).
3. The four Teachers' groups' ability to identify the words which the Students found most difficult words was compared (ANOVA).

Results and Discussion

The correlations between the four Teachers' groups' selections and the words which were unknown to the Students are shown in Table 1.

Table 1

Correlations Between Teachers' Selections

and Students' Unknown Words

	Condition	
Group	Words in isolation	Words in context
NNS Expert	.5061	.4564
NNS Novice	.5889	.5199
NS Expert	.2002	.2298
NS Novice	.0487	.0796

Basically, the correlations between the two non-native speaker Teachers groups' selections are high, both for words in isolation and words in context, whereas both of the native speaker Teachers groups show low correlations with both tests. For all of the teachers it appears to be more difficult to predict the words which will cause learners difficulties when they appear in context than when they appear in isolation. The group which is most successful in predicting the Students' vocabulary needs is the group of Novice non-native speakers (NNSN), while the group which is least successful is the group of Novice native-speakers (NNS). The correlation results suggest that teachers who speak the learners' L1 have a distinct advantage when it comes to identifying vocabulary needs. Even the Expert native speaker teachers (NSE) made a poor show of identifying the vocabulary difficulties, which suggests that their experience and training had not had a major effect on their ability to identify the learners' difficulties in this particular text. In fact, it might be argued that the NNSN group were at an advantage over their NNSE colleagues inasmuch as they had not been influenced by any linguistic or educational theory which might have interfered with their selection of the difficult vocabulary items.

Interesting as these results are, they do not really tell us how the individual members of each group performed. For example, we cannot really tell whether all of the native speakers were weak at identifying the vocabulary difficulties or whether some were good and others were bad. In order to get a more detailed picture of the behaviour within each group, the teachers' selections were analyzed in more detail. It was noted that out of the 40 words in the second (contextualised) vocabulary test, eight words were known by only 15% of the students. There were also eight words which 85% of the students got right. Ideally, the teachers should have selected the eight "hard" words and completely ignored the eight "easy" words. To find out how well each teacher performed in identifying the "hard" words and in avoiding the "easy" words, each teacher's selection was examined and given two scores out of eight: (a) the number of "hard" words selected and, (b) the number of "easy" words selected. The scores were then used to prepare scatterplots for the four groups. These are shown in Figures 1 to 4.

Figure 1

Teachers' selections of "hard" and "easy" words

(Group NNSE)

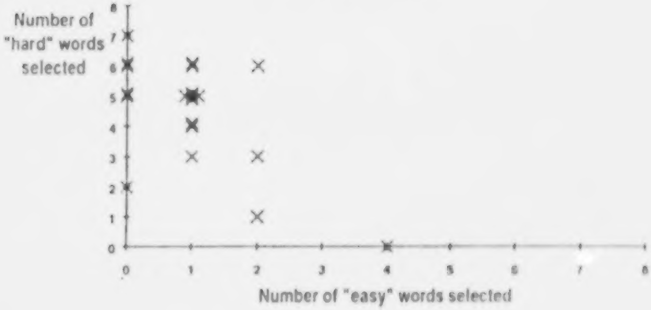


Figure 2

Teachers' selections of "hard" and "easy" words

(Group NSE)

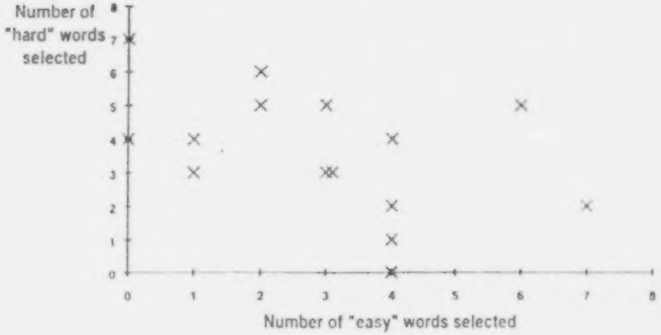
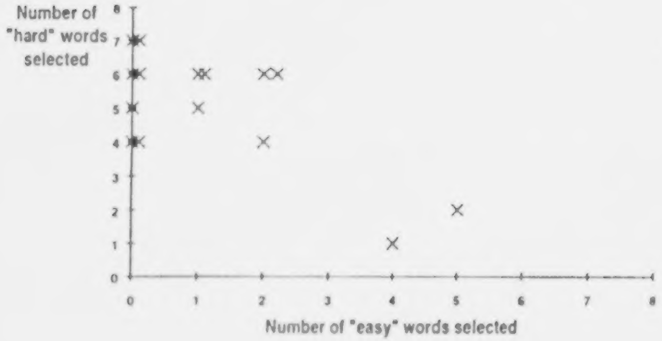


Figure 3

Teachers' selections of "hard" and "easy" words

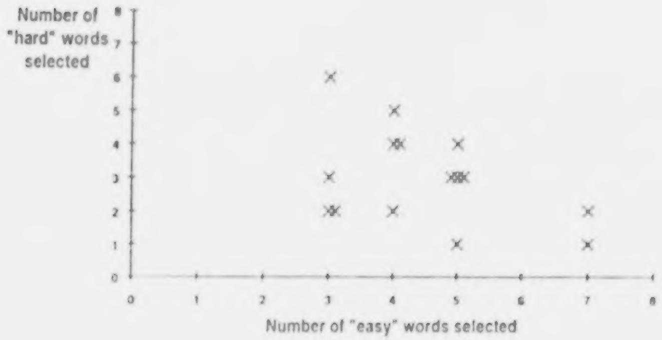
(Group NNSN)



Figire 4

Teachers' selections of "hard and "easy" words

(Group NSN)



The scatterplots provide a visual impression of how the individual members of each group performed. Good predictors can be found in the top left area of each figure, since most of their selections are from the hard words. Poor predictors can be found on the bottom right, with the majority of their selections based on words which most of the learners either already knew or could work out for themselves. The pictures suggest that both non-native speakers groups, NNSE and NNSN, tend to perform consistently well on the task, with two exceptions in each group. Although the majority make good selections, two individuals in each group fail to identify more than two "hard' words and include "easy" words in their selections. As far as the native speakers are concerned, the pattern of choice is less systematic among the Experts (NSE), with some teachers focusing successfully on "hard" words and ignoring most of the "easy" ones, while others fail to identify more than a couple of "hard" words and include several "easy" words in their selection. As for the Novices among the native speakers (NSN) a more systematic pattern can be detected. However, it is systematic in its preference for "easy" over "hard" words.

In order to establish whether the differences within each group were significant, the scores on the "hard" and "easy" words were used to calculate the phi coefficient, using the matrix in Figure 5.

Figure 5

Data table for computing phi coefficient

	"Hard" Words	
"Easy" Words	High Scorers	Low Scorers
High Scorers		
Low Scorers		

The results of the calculation are shown in Table 2.

Table 2

Phi coefficient per teacher group

NNS Expert	.58
NNS Novice	.0714
NS Expert	.69
NS Novice	.087

Significant results are obtained for both Expert groups, indicating that differences in individual performances are significant. Further research is required in order to look in more detail at the poor performers in these groups, to establish why such so many Experts found it difficult to focus on the students' actual vocabulary difficulties. The lack of significant phi coefficients for the two Novice groups suggests, on the one hand, that the non-native speakers are homogeneous as a group of good predictors and, on the other, that the native speakers are homogeneous as a group of poor predictors.

The four groups' scores on the "hard" words and on overall vocabulary judgment (calculated by subtracting each teacher's "easy" word score from the "hard" word score) were then compared by means of a one-way ANOVA. The results are shown in Tables 3 and 4.

Table 3

One-way ANOVA for "Hard" Word Identification

Source	df	SS	MS	F	p<
Between groups	3	128.3795	42.7932	19.6171	.000
Within groups	61	133.0667	2.1814		

Table 4

One-way ANOVA for overall vocabulary judgement

Source	df	SS	MS	F	p<
Between groups	3	293.6449	97.8816	12.3768	.000
Within groups	61	482.4167	7.9085		

While significant differences emerge in both measures, they are higher for the "hard" words than for overall judgement. In order to establish whether significant differences exist between particular pairs, Tukey's Multiple Range Test was run using the Overall Vocabulary Judgement scores. The results are shown in Table 5.

Table 5

Tukey's Multiple Range Test for Overall Vocabulary Judgement

Differences between groups by Language		
Native Speaker	Non-native speaker	
Expert	Expert	all significant (p<.05)
Novice	Novice	
Expert	Novice	
Novice	Expert	
Differences between groups by Expertise		
Expert	Novice	
Native speaker	Native speaker	no significant differences (p<.05)
Non-native speaker	Non-native speaker	

Significant differences emerge between all groups when compared by language, regardless of teaching expertise. By contrast, no significant differences emerge when comparing teachers by expertise within the same language.

Conclusion

The results of the study suggest that teachers who are native speakers of their students' L1 are at a distinct advantage when identifying their learners' vocabulary needs in connection with reading texts. While teaching expertise can improve non-native speaker teachers' ability in this direction, it can actually obscure the judgements of non-native speakers by interfering with their more intuitive judgements about vocabulary difficulty.

References

Berliner, D.C. (1989). Implications of studies of expertise in pedagogy for teacher education and evaluation. In the proceedings of the 1988 ETS Invitation Conference, *New Directions for Teacher Assessment*, Princeton, NJ: Educational Testing Service.

Berliner, D.C. (in press). Some characteristics of experts in the pedagogical domain. In F.K. Oser, A. Dick & J.L. Patry (eds.). *Effective and Responsible Teaching: the New Synthesis*, CA: Jossey-Bass.

Brutten, S.R. (1981). An analysis of student and teacher indications of vocabulary difficulty. *RELC Journal*, 12:66-71.

Davies, A. (1991) *The Native Speaker in Applied Linguistics*, Edinburgh: Edinburgh University Press.

Li, X. (1988). Effects of contextual clues on inferring and remembering meanings of new words. *Applied Linguistics*, 9:402-413.

Lui, N. & Nation, I.S.P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, 16:33-42.

McNeill, A. (1992). Helping ESL learners to read: teacher and student perceptions of vocabulary difficulties. In N. Bird, J. Harris (eds.) *Quilt and Quill: Achieving and Maintaining Quality in Language Teaching and Learning*, pp. 150-162, Hong Kong: Government Printer.

McNeill, A. (1993) New language, content and context: some sources of confusion for ESL teachers. In N. Bird, J. Harris, & M. Ingham (eds.) *Language and Content*, pp. 462-470, Hong Kong: Government Printer.

Mondria, J.-A. & Wit-de-Boer, M. (1991). The effects of contextual richness on the guessability and the retention of words in a foreign language, *Applied Linguistics*, 12(2).

Perkins, K. & Brutten, S.R. (1983). The effects of word frequency and contextual richness on ESL students' word identification abilities. *Journal of Research in Reading*, 6(2):119-128.

Schouten van Parreren, M.C. (1985). *Woorden leren in het vreemdetalenonderwijs.* Apeldorn: Van Walraven.

Appendix 1

Reading text used in the study

The Sword That Can Heal

While military scientists test lasers against satellites, surgeons use them as miraculously accurate scalpels. They can even be used to detonate hydrogen bombs. The beam can be focused to spot one fiftieth the size of a human hair; yet its intensity is enough to kill cancer cells or drill through the most delicate bones.

More than a decade ago, eye surgeons realised that they could use the laser's beam to seal individually, the microscopic blood vessels in the retina. The beam is so fine that only the target is heated. Now its pin-point blasting power has been turned to destroying cancer cells and reducing birthmarks. For cancer treatment, the diseased cells must be killed while their healthy neighbours are left unharmed. Where the cancer can be directly and accurately attacked, laser treatment does well: early cancer of the cervix and skin cancers have been widely and successfully treated. This type of cancer is not very easy to reach. For cancers that are less accessible, there is a new and potentially valuable technique in which the patient is injected with a chemical that then attaches itself preferentially to cancer cells. When the laser strikes the chemical, it releases a form of oxygen that kills these cells.

The marvellous accuracy of the surgical laser can be increased by sending the beam along fibres of glass far finer than the human hair. The "optical fibres" carry it around corners and direct it precisely at a tiny area; so little of the beam spills from the glass that there is no risk of damaging healthy cells.This technique is particularly useful in ear surgery.

Furthermore, the laser beam can also remove bone, and so it is invaluable in ear surgery. The sounds we hear are carried from the eardrum to the nerves of the ear by a delicate set of pivoting bones which sometimes solidify, causing deafness. A laser beam vaporises the bone without touching any of the surrounding tissue. The beam is diffused to avoid scarring and the mark becomes inconspicuous. This accuracy in targeting makes the laser a useful tool for the dentist also - a nerve can be reached through a hole drilled in the enamel.

Birthmarks, once almost untreatable, are a mass of blood vessels and, being red, they absorb the laser beam strongly. It seals them so that the mark becomes less conspicuous. The normal cells of the skin's surface, which don't absorb much of the laser beam, act in the healing and help to conceal the mark. The beam can cut with a precision that no scalpel could achieve. The operation can transform the lives of people who were previously doomed to a lifetime of cosmetic surgery.

Though this application is widely used in America, there are in Britain only two hospitals offering the treatment, and one feels bound to warn patients that success is not certain. However, some ten new centres will soon be opened.Britain, though, is one of the leaders in the laser treatment of bleeding peptic ulcers and this, combined with new medicines can mean ulcer treatment without conventional surgery. The laser is now being used to treat all kinds of illnesses in this country.

(Tony Osman in The Sunday Times Colour Magazine, reproduced in S. Greenall and M. Swan. *Effective Reading,* Cambridge: Cambridge University Press, 1986)

USING CONCURRENT VERBAL REPORTS AS TEACHING TOOLS IN LANGUAGE TEACHER EDUCATION

Margaret Falvey

Introduction

In an age of rapid technological advance, higher order reading skills are now essential as a means of access to new information. As a result, it is even more important than before that language teachers should be aware of the many and complex processes involved in reading and that their approach to the teaching of reading should be based on a firm theoretical underpinning.

As knowledge about the nature of the reading process has increased, however, so has the content-load in this aspect of language teacher education programmes. Further increases in this content-load have resulted from, for example, research into the nature of learning. As the volume of theory grows so does the risk that it will be rejected by student-teachers as having no relevance for classroom practice. If this happens the theory will have no influence on their teaching post-course even though they may pay lip-service to the theory during the course in order to obtain their qualification. Achieving the right balance between theory and practice is therefore important on initial teacher education programmes if the programmes seek, to:

1. have long-term effects on teacher thinking and teacher practice and
2. provide an appropriate foundation for future in-service teacher-education programmes.

This means that the approach to theory must be selective.

Approaches to the role of theory in education courses

One approach, no longer regarded as acceptable, was to select one theoretical position, either the most recent or merely that preferred by the teacher educator, present it as 'The Way' and concentrate on giving student-teachers practice in applying it in the classroom. Student-teachers completed such courses unaware of alternative views and lacking skills as critical consumers of research. As a result, they were ill-equipped to understand, let alone apply, theoretical views which emerged after their professional careers had commenced.

Another approach has been to provide student-teachers with a broad introduction to theory in the relevant area and then require them to do a lot of reading, culminating in written assignments to demonstrate their understanding.

Unfortunately, student-teachers are often reluctant to read research papers. They are put off by the statistics, the language and particularly by the de-humanising way in which some researchers treat their human subjects and write about them. They are also sceptical about the value of data obtained under laboratory conditions rather than in every day classrooms. Student-teachers are not, of course, alone in this view; the development of alternatives to the logical-positivist approach, to the experimental approach and to some aspects of quantitative research have come about because researchers into education, amongst others, hold similar views.

An alternative approach

An alternative approach to those described above is preferred by this researcher. This is a 'grounded theory' approach in which the theory and the research design are developed from and informed by data (Glaser and Strauss 1967, Hammersly and Atkinson 1983). In this case the student-teachers first encounter research data and then explore and discuss them in order to create a shared context and shared information for discussion of different theoretical perspectives. In the case of research into reading, concurrent verbal reports would appear to offer ideal data for this purpose for two reasons. First the data are very clearly generated by the individual subjects and are full of 'human touches.' Secondly concurrent verbal reports reveal not only the cognitive processes involved in reading but can also, under certain conditions, reveal the sort of problems that readers, particularly student readers, may encounter when trying to understand text. Published research is not the best source of data in this case, however, because research papers do not usually offer data in the form needed for this purpose. They tend to report findings and results rather than raw data and presenting these to student-teachers has the drawbacks already discussed above.

Finding suitable data

If, despite the problems discussed above, the teacher educator feels that student-teachers would benefit from exploring data obtained through verbal reports some alternative way of creating access to these data must be devised.

This researcher has observed that subjects often become intrigued by their own cognitive processes during experiments using verbal reports. This seems to occur in much the same way as it does with clients receiving counselling who are, for example, fascinated to discover that their expectations of other people are based on beliefs they were not even aware they held.

The initial stages of the present study were based on the above observations. It was anticipated that student-teachers would be intrigued by their own cognitive processes and would therefore be interested in exploring and analyzing data which revealed them. As a result the student-teachers themselves became both the subjects and the researchers in the present study.

It was predicted that the data would have high face-validity because the student-teachers, having themselves been the subjects, would know that the data were not the outcome of researcher-manipulation of subjects. It was also anticipated that the experience of being first the subjects, and subsequently the analysts of the data, would give them access to some of the formal and content schemata needed to understand lectures and discussions about research into reading. The possibility that some of them might even store this new information well enough to enable them to read research papers in this area with an initial degree of expertise and ease was regarded as an extra bonus which might encourage voluntary reading in this area, in the future, even if not during the course.

In the section below concurrent verbal reports are defined in the context of this study.

Concurrent verbal reports

In this paper concurrent verbal reports are defined as 'subjects' verbalizations elicited while-reading.' The reports in this study were elicited by asking the subjects to act as if they were alone in a room, speaking to themselves. They were asked to 'say aloud everything you say to yourself silently' (Ericsson and Simon 1993). They were asked **not** to explain their thoughts but merely to say them out loud.

The lack of coherence reported in concurrent verbal reports elicited in psychological studies is not so evident in the data reported here. Possible explanations for this are offered in the section *Problem-solving Tasks vs. Reading Tasks*.

Selecting a reading text

The selection of the text for this study was restricted by the context of implementation. The text had to have high face-validity, i.e. it had to be the sort of text that teachers need to read. It also had to be suitable for student-readers so that the complexity of the reading processes identified through the verbal reports could not be dismissed or attributed to the fact that the text required an expert reader. At the same time the text had to be authentic in the sense that it should not be the type of text commonly found in EFL textbooks which are specially written to give learners practice with particular language structures or vocabulary items. Finally it was necessary that the text should require complex processing on the part of the reader and should also present some reading problems for the readers in this study so that the resulting verbal reports would illustrate both.

The text which was used in this study meets the above criteria. It is discussed in more detail in the section *The Text and The Readers*.

Planning a learning experience vs. a research exercise

The study reported here was planned primarily as a learning experience and not a research exercise. It was not, therefore, considered desirable to try to replicate laboratory conditions in which each subject is isolated even from the researcher. The purpose was to provide a reasonably natural and shared 'reading' experience which the student-teachers could discuss while the experience was still fresh in their minds. As a result there was no training for the subjects and the usual classroom practice of working in groups was adhered to.

The student-teachers were told that the purpose of the session was to explore the processes involved in reading; to find out 'what goes on in people's heads while they read.' Each group of four chose a 'reader', or subject, who would provide a concurrent verbal report while 'reading' the passage. The 'readers' were given the instructions stated earlier in this paper. Each group also selected a person known as the 'text' who would read the selected text aloud, repeating parts of the text and spelling words aloud if asked by the 'reader', but not interacting with the 'reader' in any other way. The 'text' would not, for example, respond to questions or remarks made by the reader.

Transcripts made from audio-recordings of the verbal reports were not regarded as a satisfactory means of access to the data since it takes time to produce them and the delay could reduce both the impact of the experience and the student-teachers memory of it. Instead, instantly available verbal reports were generated by having two 'Observers' attached to each 'reader,' or subject, whose tasks were:

1. to make complete transcriptions of everything said by the subject and

2. to prompt the subject to keep talking when necessary.

The data reported in this paper are, however, based on full transcripts made from audio-recordings of the entire proceedings in the classroom.

Analysis of data

In this learning experience the student-teachers, who were both subjects and researchers, explored the initial data themselves. Since the learning experience was based on a 'grounded theory' approach there was no attempt to provide them with methods of analysis. Their initial analysis, which was conducted through group discussion of the data, was recorded and that data form part of the long-term study of which only the first stage is reported here.

The writer's own initial analysis of the data was conducted by moving around the classroom and taking notes during the elicitation of the verbal reports. This analysis was also exploratory but had a slightly different agenda in that the purpose was to identify and label specific processes employed by the subjects to facilitate future discussion of the nature of the processes involved in reading. Although the data reported in this paper are based on transcripts of audio-recordings of the concurrent verbal reports, the commentary on the data reflects the above approach, i.e. it is exploratory and represents a 'hands-on' analysis.

The following section briefly outlines the differences between the nature of the tasks set in psychological research and those set in language education research into cognitive processes.

Problem-solving tasks vs. reading tasks

Research into cognitive processes in areas such as problem-solving employ tasks which are short and 'well-defined.' Such conditions are deemed necessary for the study of reproducible performance. A task is 'well-defined' if:

1. it has a clear focus, e.g. '48 x 27' and

2. 'only a limited number of possible sequences of thoughts will generate the correct answer efficiently' (Ericsson and Simon 1993).

The instructions for such tasks, which may involve written text, will also fulfil these conditions. The cognitive processes in the resulting verbal reports will therefore have been employed mainly to solve the problem and not to understand the instructions.

In language education research into text comprehension, however, the statement of the task and the problem itself tend to be one and the same thing, i.e. the text. The text may, for example, be a text that student-readers need to read for study purposes but find difficult to comprehend. The performance of 'expert' and 'novice' readers may be compared to identify the strategies employed by the former which could usefully be taught to student-readers (Lundeberg 1987). The reading task is therefore unlikely to be either short or well-defined as described above and comprehension will probably require interaction between many complex cognitive processes. These interactions will vary from one reader to another even in the case of 'expert' readers who are Idealised Target Readers (ITR's) (Falvey 1993), because of differences in, for example, their prior formal and content schemata. These differences will probably be greater in the case of 'novice' readers, especially if they are not even 'Marginally Related Readers' (MRR's) (Sinclair 1993). Because of these differences the resulting verbal reports can be expected to vary considerably. Reading texts, therefore, usually present 'ill-defined tasks,' which, according to Stratman and Hamp-Lyons (1994), involve one or both of the following features:

1. subjects must specify partly or completely their own goals and

2. subjects may generate many equally satisfactory solutions' (Stratman and Hamp-Lyons 1994).

The text used in this study presents an 'ill-defined task' and, since the subjects were neither ITR's nor even MRR's, individual subjects' efforts at comprehension were expected to generate a wide range of different cognitive processes. It was therefore anticipated that the verbal reports would offer evidence of a number of the different processes involved in reading, thus fulfilling the goals of the learning experience.

It was also anticipated that the verbal reports would exhibit more coherence than is the case in problem-solving studies because the selected text:

1. encourages social interaction between itself and the reader by using the first person, 'I' and 'me' and addressing the reader as 'you';

2. would require the subjects in this study to employ considerable controlled and off-line processing in their efforts to extract its meaning.

The text and the readers

The selected text is authentic in the sense that it was published in a 'fun-reading' book. The ITR's were Primary Four pupils in Eastern Nigeria who were studying English as a second language. The text is a riddle in the form of a rhyme with 6 stanzas and 4 lines to a stanza. It displays many characteristics of the genre of riddles in that the macro-structure is signalled in the title, 'What Am I?' indicating that:

- the reader's task is one of problem-solving, i.e. to guess the answer to the question in the title.

- there will be clues to help him

- these clues will probably not be explicit

- all the information given will be 'true' but many alternative meanings may need to be considered

- the problem-solving process may involve lateral thinking

Because this text is longer than the average riddle, the reader receives more clues and has more opportunities to predict and confirm predictions than is the case with most riddles or in the case of the other riddles of which the ITR's had had experience. The structure of the text is briefly outlined below:

Stanza 1: Non-explicit reference to plots of traditional African stories about the tortoise, most of which involve the cracking of its shell.

Stanza 2: Explicit references to 'tortoise stories.'

Stanzas 3-6: Explicit reference to three of these stories, the first including implicit reference and the other two making explicit reference to the cracking of the tortoise's shell.

The ITR's could be expected to predict the correct answer during the first stanza, using the remaining five stanzas to confirm their ideas. Alternatively there are four more opportunities which could be confirmed by backwards reference facilitated by the overtness of the text structure which helps the reader to remember preceding lines. Because of their previous experience of the book from which the text was selected the ITR's would expect to:

- understand the text without help from the teacher,
- enjoy reading it,
- interact with the text.

These expectations could be considered to be justified because:

- all language items, structures and vocabulary were familiar;
- none of the texts previously read would have been used to teach new language;
- the formal schema of riddles was already familiar through exposure to shorter riddles;
- the activities associated with the texts were all 'fun' activities, i.e. miming or acting out the stories or making up their own riddles;
- they possessed the necessary content schema through listening to and reading traditional African stories.

The ITR's could therefore be expected to employ mainly automatic processing while reading the text. Although the readers in this study were adults, highly educated and generally skilled readers they were, by comparison with the ITR's, operating at a disadvantage since they:

- were unlikely to have the necessary content schema;

- might lack the formal schema of riddles;
- were told that the text was taken from a Primary Four 'fun' reading book but their own learning-to-read experiences had not included reading for pleasure.

It was therefore predicted that:

- even those who had the necessary formal schema might fail to activate it on reading the title;
- all of them would have difficulty in processing the text successfully, i.e. in predicting and then confirming the one correct answer to the question 'What Am I?' with any degree of certainty.

The text was therefore considered ideal for the purposes of this learning experience and the resulting verbal reports were expected to yield a rich variety of both the processes involved in reading, the types of problems that learner readers may encounter and the strategies that skilled readers use to deal with these problems.

The Concurrent Verbal Reports

The concurrent verbal reports cited here were elicited during the first of each subject's three readings of the text. S.1-S.3. are the subjects. Text (T) is the group member who read the text, a riddle, aloud. Line numbers, given in brackets at the end of each line of the text, are inserted in the verbal reports to indicate the point in the text at which comments were made. If there is no verbalisation cited from a subject it means the text continued to be read aloud until that subject did verbalize.

Although the concurrent verbal reports shown below were elicited separately, Examples 1-6 (in response to lines 1-16 of the text) are presented together to avoid unnecessary repetition of the text.

These reports were collected primarily as part of a learning experience in which student-teachers first provided and then explored the data themselves. The commentaries reflect the 'hands-on' analysis of a teacher-educator looking for

a. evidence of the processes the subjects employed during their first reading of the text and

b. differences in the processes used by different subjects. Not all processes are commented on. Those selected are those deemed to offer:

1. easily identifiable examples of the processes,

2. examples which will encourage further discussion of the non-verbalised but inferrable processes

3. those which will encourage discussion of the multiple and complex interactions taking place between processes.

The processes identified include the following; the activation of formal and content schema (FS and CS), automatic and controlled processing (AP and CP) Schneider 1985), on-line and off-line processing (OLP and OFP), processing at the levels of macro-structure and micro-structure (MOS and MIS), the psycho-linguistic guessing game (PLG) (Goodman 1968).

Example 1. Title and Lines 1-4.

T: What Am I?
I carry my house (1)
Around on my back (2)
So I walk very slowly (3)
Because it might crack. (4)

S.2. (Title) * What am I? OK. Go on.
S.3 (2) - Em, em, it is a snail, I think. Er, is it, is it? Well, it is a snail. Go on. (laughs)
S.1 (4) - It is quite difficult to read the sentence. Let me read it again.

T: Repeats title and lines 1-4
S.1. (4) - I think I can get a rough idea of that. I think I can get a rough idea of that sentence. A boy carrying a bag. OK. Lets' go on.
S.2. (4) - So you are a little animal
S.3. (4) - I walk every day but I never thought of cracking myself. But it is a problem of a snail, a poor snail and --- go on.

Although only S.2 comments on the title, by repeating it, both S.2 and S.3 appear to have activated the formal schema (FS) of riddles. S.3 predicts 'snail' (2) and S.2. appears to both predict and confirm with "So you are a little animal" (4). The speed of response suggests that both were employing automatic processing (AP). They were also employing off-line processing (OFP) and using their prior content schema (CS) but were retrieving different information and applying it in different ways. S.3 appears to find a contradiction (4) but then confirms 'snail' with what may be another prediction. This subject seems to be formulating ideas about the macro-structure (MOS) by anticipating that the overall purpose of the text is to tell a sad story about a snail, which would not be surprising in a primary school reading book. It seems that neither S.2 nor S.3 possess the necessary CS. The use of 'you' suggests that S.2 is interacting with the text in response to 'I.'

S.1. seems to be in trouble, the FS for riddles does not seem to have been activated. This may be because, knowing this text is a primary school reading text, she does not expect something frivolous. She interprets the 'I' as:

a) human and

b) child.

Example 2 (lines 5-8):

T: Don't you believe me? (5)
It's happened before. (6)
You remember the stories (7)
You've heard them I'm sure (8)

S.3 (7) - (interrupts) What. what, stop. It happened to the snail before, that means the snail cracked himself before. And then, go on.
S.1 (8) - Who is **that** you? Maybe I'll read it further to find out who is you.
S.2 (8) - What kind of story?

There is a distinct change in the text in these 4 lines: the reader is directly addressed as 'you' and is presented with a challenge, followed by a statement and then a reference to CS which is clearly assumed to be shared. Since it is not shared by these readers, one would anticipate some sort of reaction.

S.3 appears to respond, employing AP and OFP. She interrupts the text, but seems to drop her query about the stories and to focus on the implications for the MOS. She re-formulates her thinking to allow cracking shells to fit with her prediction of a sad story. S.2 does respond (AP and OFP) but does not follow up. S.1 is provoked but does not relate to the 'you' interactively. Instead she wonders who 'you' is. This subject seems to assume that the text itself will ultimately make everything clear and that information in the text will be 'literally' true. She does not seem prepared to make inferences even through controlled processing (CP). It seems that both 'you' and 'I' are assumed to be human.

Example 3 (Lines 9-12)

T: The bird took my feathers (9)
So I couldn't fly (10)
The ground's very hard (11)
When you fall from the sky (12)

S.1 (10) - Quite abstract, this boy's thinking. The bird took my feather. It seems that he imagines himself as a bird. Let's read further.
S.2 (10) - Oh, so you are not a bird.
S.3 (10) - Hold on, the birds, would you repeat that again?

T: Repeats lines 9 and 10

S.3 (10) - The bird took his feathers, so it is not a snail. Go on.

S.2. uses AP and OFP to rule out one unpredicted possibility and continues to address the text as 'you.' S.3 is also using AP. There is a delayed reaction to the birds and then a different inference which rules out her original prediction 'snail.' S.1. seems surprised by the nature of the text, by the level of abstraction, so one would expect some re-formulation of her expectations about the text. She makes a tentative inference using AP and OFP and then moves to CP. Perhaps this subject has assumed that in a primary level text all personal pronouns, with the exception of 'it', will refer to humans. As before, her solution is to proceed with reading rather than resort to CP or OFP.

Example 4.

T: The ground's very hard (11)
When you fall from the sky (12)

S.1 (12) - It seems that he really thinks that he is a bird.
S.2 (12) - What kind of animal you are? (laugh)
S.3 (12) - Hold on, how do you spell the word 'ground'?

T: * G_R_O_U_N_D

S.3. * Oh, the ground's very hard, OK go on.

T: Repeats line 12

S.3. * the ground's very hard when you... go on.

S.1 uses AP and OFP and finds confirmation of her inference but considers it unreal or unlikely. S.2 is still interacting with the text, using AP and OFP. She is sure that 'I' is an animal but finds nothing on which to build a further prediction. S.3 thinks she has misheard and asks for a word to be spelled out loud. Her next comment suggests that while lines 11 and 12 then make literal sense, she is, nevertheless, still trying to make sense of the last 4 lines in the context of her previous interpretation. She may be trying to rationalise her first prediction of 'snail' to make it fit with the new information that has come on-line or she may be trying to re-interpret, or 'tune' (Rumelhart and Norman 1978), the new information to make it fit with her predicatior of 'snail.'

Example 5 (lines 13-14).

T: The monkey tricked me (13)
With his tail I fell (14)

S.1 * Hm, monkey? Oh, er, that, before he said he is a bird. But now why is there come out a monkey? I will read further to find out the answer.
S.3 * (Interrupts) Could you repeat that again?

T: Repeats lines 13 and 14
S.3 * With his tail? How do you spell the word 'fell'?

T: F_E_L_L_
S.3 * The monkey tricked me with the tail I fell what?

T: Repeats lines 13 & 14
S.3 * That's the end? (laughs)

S.3 seems to be experiencing learner-confound. She requests repetition and the spelling of words, for example. She may, however be re-processing, or trying to tune, the textual information given so far. She seems to react against the partial text, caused by her own interruption.

S.1 is using CP and OFP but she is still in a very literal world and still relying on the text to reveal all.

S.2 is silent.

Example 6.

T: Down through the branches (15)
Another cracked shell! (16)
S.2 * You live on the tree?
S.3 * Can you repeat that sentence, start from, read from 'the birds took my feathers--'

T: Repeats lines 7-14.
S.3 * That's the end? OK, go on.

S.2 seems to be considering different types of animal. She infers the habitat of 'I.' Good examples of AP and OFP. S.3 seems to be reviewing all the textual information with which she has experienced a problem. She specifically asks the text to repeat from "the birds took my feathers..." Interesting that she remembers the exact words which suggests that the linguistic data is easy to handle but the semantic data is not. She doesn't seem very happy by the end of the repetition, however, as though she expected a 'click of comprehension' (McLeod and McLaughlin 1986) which didn't come. This subject seems to have a grasp of the MOS and is focusing on those elements in which she expects to find a solution. She seems to be quite clear about the roles of lines 1 to 8 in the MOS and also that lines 9 to 14 merit attention.

S.1 makes no comment here. In fact, after a pause the text went straight on to the next two lines (17 and 18). S.1 didn't interrupt or comment until the end of line 18 (see Example 9).

Subjects' verbal reports in response to lines 17-24, see below, are presented separately in Examples 7-9 for ease of reference.

Example 7 (lines 17-24).

(S.3. only).

T:
Trying to save money (17)
By making some oil (18)
I climbed on the pot (19)
As it started to boil (20)
But the pot was too hot (21)
And I started to fry (22)
So I jumped to the ground (23)
Oh! I wish I could fly! (24)

S.3 (23) - (Interrupts after 'jumped' in l.23)
Er, hold on, the birds took his feathers, something, something that has feathers, and then being fell, it couldn't fly because it doesn't have feathers, then, the monkey (laughs), the monkey tricked it with its tail, then, er, then he try to sell, what, make money? Read from the tail, that part.

T: Repeats lines 13-21.
S.3 * (Interrupts after 'pot' in l.19).
* Trying to make, hold on, er, I guess oil makes a lot of money. And er well, you can use oil for many things, right? So, of course, it makes money. I hope I can have oil. Maybe. Large piece of land, lots of oil underground, then I can make money from that, well, go on. Trying to the pot...what? Go on.

T: Repeats lines 17-21.
T: Reads lines 22-24.
S.3. (24) - Well, well, if the, something that couldn't fly because the birds took his feathers, so it tries to fly but it couldn't, it tries to jump, er--, OK, starting from I climbed on the pot
T: Repeats lines 19-24
S.3. * Er, go on. That's the end? Em, it jump to the ground and then, well starting from 'I climb on the pot' again

T: Repeats lines 19-20 again
S.3. * (Interrupts) I jump to the ground, I wish I could fly, because the pot was too. because the pot was too, well, the pot was boiling, of course it was hot, so because it's hot, it wants to escape from it, so it jump. Where does it jump? Jump from the pot, jump from the pot. So this thing it crack its shell, somebody, the bird takes his feathers, he couldn't fly, jump? what is it? I cannot make any meaning from it (laughs) I can't think of anything. Well, start from the beginning again.

(End of Subject 3's first reading)

S.3 is still working on elements of the MOS, trying to re-organise the textually new information in lines 9-14. Although her first comment looks like CP, one can infer a lot of AP, involving interaction between the subject's own prior schema and the textually new information. Her apparent errors, the notion of 'selling oil' and 'making money' strongly indicate AP. In her second comment this subject is determined to 'make sense out of nonsense.' She is involved in extensive OFP and in complex interaction between her prior CS and textually new information. The comments about lots of land and oil underground suggest she has animal habitats in mind and then it is clear that the 'pot' doesn't fit into this new schema or with the idea of an animal.

By the end of the last line the 'I' has become a 'something' and S.3 has identified some of 'its' characteristics. She seems to infer that: with feathers it could fly; then without feathers it couldn't fly; it jumps or it tries to jump. Then, however, 'the pot' causes a problem again. After the requested repetition of lines 19-24 the subject is surprised that this is the end of the text. She then goes back to the problems with 'jumping' and 'the pot.' After another repetition, this time of lines 19-20 only, she interrupts, summarises the contradictions and explains some of them to her own satisfaction. She then reviews the important characteristics until she gives up and asks for another reading. Once again this subject demonstrates complex interaction between her perceptions of the MOS and MIS, her own prior CS and textually new, but already processed information. Most of this part of her report involves CP and OFP.

Example 8 (lines 17-24).

(S.2. only).
S.2 (19) * You are making money?
T: Yes (should not have responded)
S.2 * I don't understand... he make money.... as you said. Could you repeat this sentence?

T: repeats lines 17 and 18.
S.2 * save some money by making oil. Go on.

T: Reads lines 19-20.
S.2 * Pot. How to spell pot?

T: P_O_T
S.2 * Ah-ha. Then what happened to you? Is it very hard? And are you an ant or something else? Go on, go on.

T: Reads lines 21-23.
S.2 * (Interrupts after "I" 1.23)
Started to fly? You have read... you just say you haven't got your feathers. Go on, go on.

T: Reads lines 23-24.
S.2 * But you can't fly now. I still don't have any idea what kind of animal you are--- you live on the tree, you have no feathers. How can you fly? go on (laughs)----Would you please go on?

T: That is the end.
S.2 * Could you repeat all the passage again?

(End of Subject 2's first reading)

S.2 has a problem with the idea of an animal making money. Her interaction with the 'I' in the text is confirmed by her confusion when the text actually answered her question. This subject employs automatic processing to retrieve, from her prior content schema, an animal (insect) that is likely to 'climb on a pot', an ant, but isn't entirely happy with this choice. She then mishears 'fry' and hears 'fly' instead. The text definitely said 'fry' however. This 'error' does not seem to be the outcome of problems with 'l' and 'r' sounds. In fact it is quite a logical 'error' in the context. The subject's exposure to the textually new information and to the linguistic forms used so far in the text would make 'fly' an expected action whereas the only contextual preparation for 'fry' is the word 'pot.' Since this subject perceives 'I' as an ant, she is unlikely to consider 'I' as part of the food in the pot and is more likely to regard the ant as an undesirable extra. Why then should she expect him to be fried? Much more likely that the ant 'started (tried) to fly' in order to escape from the hot pot. She reacts against the contradiction using OFP but does not pursue it once she hears line 24. This subject then reviews certain aspects of the textually new information but re-confirms the contradictions she had identified earlier. Like S.3, she demonstrates complex interactions between her own prior CS and recently processed textual information using CP and OFP. She too assumes that there will be more text.

Example 9 (Lines 17-24).

(S.1 only)

T: Reads lines 17-18.
S.1 * Trying to save money by making----? Oh, maybe. Repeat that sentence.

T: Repeats lines 17 & 18.
S.1 * Oh, trying to save money by making some oil? Maybe I'll read further.

T: Reads lines 19-20.
S.1 * Pot? What's the spelling, the spelling of the word 'pot'?

T: P_O_T_, pot.
S.1 * As it started to boil? repeat that sentence.

T: repeats lines 17-20.
S.1 * Oh, climb on the pot, it started to boil? Is there any relationship with the money and save money then? It seems that it hasn't any relationship at all. But... let's read further then.

T: Reads lines 21-24.

S.1 * Oh, it seems that, does it mean, does it mean that he has fallen into the pot and fly again. Maybe, er, I read, er, repeat the sentence then.

T: Repeats lines 21-24.
S.1 * Oh, it seems that he is still imagine himself as a bird, just like at the beginning of the passage. I think that is----- Maybe I'll read it once again to get more idea.

(End of Subject 1's first reading)

During lines 17-20 S.1 seems to be experiencing serious learner confound and appears unable to process the text in units of meaning, focusing instead very much at word-level and forgetting, or ignoring, lines she has just heard. The continuing mention of 'relationship' seems to refer to a relationship between words, or types of words, rather than to relationships between units of meaning and suggests that she is using a strategy of 'word-matching' which is unfortunately sometimes taught as an aid to comprehension. As a result she is operating only at word-level most of the time. By the end of line 24 she moves tentatively into OFP but is still operating very literally.

Conclusions

Data from the first set of concurrent verbal reports provide evidence of AP, CP, OFP, OLP, activation of CS and FS and interaction between these and the subjects' perceptions of MOS and MIS. The reports also demonstrate the nature and causes of some of the reading problems encountered by learners. The data therefore meet the needs of the learning exercise. In addition, they create openings for the discussion of a number of issues related to the teaching and learning of reading.

The fact that none of the subjects got the 'right answer' emphasises the importance of prior CS and offers evidence that some of the problems that foreign language learners experience with reading may be precisely that, i.e. reading problems, not language problems. The data contradicts the popular assumption that the answer is always 'in the text', literally and demonstrates the importance of OFP and interaction between prior CS and textually new information. The data shows that even young readers, faced with a linguistically 'easy' text, need to employ many complex processes including, amongst others, OFP, OLP, AP and CP and to interact with the text at the level of MOS as well as MIS. This illustrates the need for young readers to learn:

1. to operate at the level of discourse instead of operating at word or sentence level only and

2. to activate their own CS and FS instead of relying on teacher interpretation (Falvey 1986).

The many differences between the three verbal reports offer rich material for discussion of the importance of recognising that different readers process text in different ways.

The reports of S.1 and S.2 yield good examples of Goodman's 'psycho-linguistic guessing game;' S.1's report demonstrates the disadvantages of not playing the game. In an educational context in which 'right answers' have more value than alternative legitimate interpretations and in which guessing is frowned upon the relative success of subjects 2 and 3, who were guessing, provides a useful contrast with the relative failure of subject 1.

Some of the apparent 'hearing' errors, or slips made by subjects when repeating text, provide evidence that skilled readers do not faithfully store the exact words they read. Instead they store the meanings that have emerged as the result of processing. This aspect of the data provides a good introduction to the issue of testing reading competence by making learners read aloud and checking on their 'accuracy.' Such discussion is further enriched by the retrospective verbal reports from the 'texts' who said they were 'unable to think about extracting meaning' because they 'had to concentrate on reading aloud.'

The subsequent classroom discussions on reading processes and reading problems were based on explorations of three sets of data all of which were collected in the same session:

1. the concurrent verbal reports cited in this paper
2. two further concurrent reports elicited during each subjects' second and third readings of the text
3. retrospective reports elicited in groups consisting of
    a. the three subjects,
    b. the three texts and
    c. the six Observers.

The data proved very accessible to the student-teachers. The role of the teacher-educator was mainly that of supplier of meta-linguistic labels.

It appears from the data described above that concurrent verbal reports are useful tools for raising student-teachers' awareness of the processes involved in reading. It is worth considering whether there are any other research tools which could be used or adapted for pedagogic purposes in teacher education.

References

Ericsson, K.A. & Simon, H.A. (1993). *Protocol Analysis: Verbal Reports as Data: A Bradford Book*, Cambridge, MA: MIT Press.

Falvey, M.A. (1986). The Reading Teacher's Dilemma. *Working Papers in Linguistics and Language Teaching*, April 1986:44-57. Hong Kong: University of Hong Kong, Language Centre.

Falvey, P. (1993). Towards a Description of Corporate Text Revision. Unpublished Ph.D. thesis, University of Birmingham.

Glaser, B., & Strauss A. (1967). *The Discovery of Grounded Theory*. Chicago, IL: Aldine.

Goodman, K. (ed.), (1968). *The psycho-linguistic nature of the reading process.* Detroit, IL: Wayne State University Press.

Hammersly, M. & Atkinson, P. (1983). *Ethnography: Principles in Practice.* London and New York: Routledge.

Lundeberg, M.A. (1987). Metacognitive aspects of reading comprehension: Studying understanding in legal case analysis. *Reading Research Quarterly,* 22(4):408-432.

McLeod, B. & McLaughlin, B. (1986). Restructuring or Automaticity? Reading in a Second Language, *Language Learning,* 30(2):109-121.

Rumelhart, D.E. & Norman, D.A. (1978). Accretion, Tuning and Restructuring: Three Modes of Learning. In J.W. Cotton & R.L. Klatzky (eds.), *Semantic Factors in Cognition*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Schneider, W. (1985). Toward a Model of Attention and the Development of Automatic Processing. In M.I. Posner & O.S.M. Marin (eds.), *Attention and Performance XI*. Hillsdale, NJ and London: Lawrence Erlbaum Associates.

Sinclair, J.McH. (1992). Writer, Text and Reader. Unpublished paper, University of Birmingham.

Stratman, J.F. & Hamp-Lyons, L. (in press). Reactivity in Concurrent Think Out Loud Protocols. In P. Smagorinsky (ed.), *Verbal Reports in the Study of Writing: Problems and Potentials.* Newbury Park, CA: Sage Publications.

Reprography by the Government Printer, Hong Kong 255021-24L-12/94

BEST COPY AVAILABLE